13th Int. Workshop on Generalized Nets
London, 29 October 2012, 8–14

Generalized net model of hierarchical neural networks

Krassimir Atanassov[1,2], Sotir Sotirov[2] and Anthony Shannon[3]

[1] Institute of Biophysics and Biomedical Engineering, Bulgarian Academy Sciences
105 "Acad. G. Bonchev" Str., Sofia–1113, Bulgaria
e-mail: krat@bas.bg

[2] "Prof. Asen Zlatarov" University
1 "Prof. Yakimov" Blvd, Burgas–8010, Bulgaria
e-mail: ssotirov@btu.bg

[3] Faculty of Engineering & IT, University of Technology, Sydney
Sydney, NSW 2007, Australia
e-mail: Anthony.Shannon@uts.edu.au

Abstract: We construct here a Generalized Net (GN) that represents the functioning and the results of the work of real processes and simultaneously – the processes of their control and optimization on the basis of different suitably chosen hierarchical neural networks solving concrete optimization procedures and using information generated in the GN.
Keywords: Control, Generalized net, Neural network, Optimization.
AMS Classification: 68Q85, 62M45.

1 Introduction

A series of many generalized net models have been constructed so far, representing the way of the work and the optimization of different types of neural networks (NN). The first one presents NNs which are learned with teacher (supervised neural networks) [1, 6–9, 13, 15].

In the process of learning with teacher, a result has to be achieved, which is known in advance thus giving self-regulated direction to the neural network. The weight coefficients are accordingly changed to achieve a fixed by the 'supervisor' quantity. After its learning, the neural network passes a test – only entry signals are submitted, without the signal which must be obtained. The concrete exit values are received on the network's exit.

Some types of these NNs introduce the work of the feedforward NNs, that was described by GNs [6, 15]. The GN-models in [7–9] describe the 'backpropagation' learning algorithm for neural networks; the GN-model in [13] presents an accelerating learning of the 'back-propagation' algorithm from [12].

The GN-models of the next type NNs, namely 'self organizing map' (SOM), were constructed in [5, 14]. The SOM is a subtype of the unsupervised artificial neural networks. It is trained using unsupervised learning to produce low dimensional representation of the training samples, while preserving the topological properties of the input space. This makes SOM reasonable for visualizing low-dimensional views of high-dimensional data, akin to multidimensional scaling. In [14], the GN-model of the work of the SOM is presented, while the GN-model in [5] describes the optimization of the SOM with time-limits.

The subsequent GN-model represents the work of another kind of unsupervised NNs, the Grossberg neural network, [17]. This NN uses a model based on the human visual system.

The GN-model in [16] represents the work of the Learning Vector Quantization (LVQ). The LVQ network uses both unsupervised and supervised learning for each classification. In LVQ network, each neuron in first layer is assigned to a class, with several neurons often assigned to the same class. This type of neural network is different than other. It combines supervised and unsupervised NN.

Here we present one generalized net model that describes hierarchical combining of six neural networks. In it we can choose different structure of the supervised neural networks.

2 The generalized net model

All definitions related to the concept of GNs are taken from [2–4, 10, 11]. The network, describing the work of the neural network learned by the ‘backpropagation’ algorithm, is shown on Figure 1.

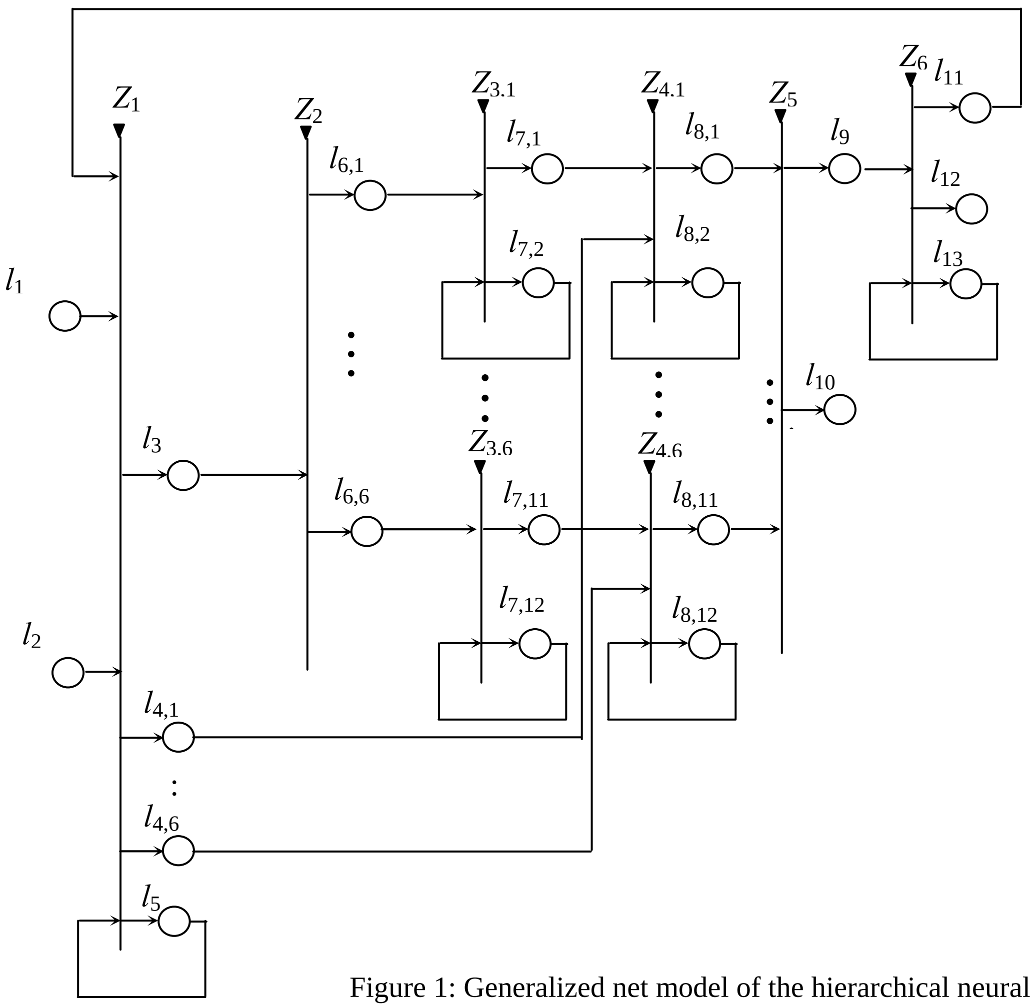

Figure 1: Generalized net model of the hierarchical neural networks

The below constructed GN-model is a reduced one. It does not have temporal components, the priorities of the transitions, places and tokens are equal, the place and arc capacities are equal to infinity.

Initially, the following tokens enter in the generalized net:

- in place l_1 – σ-token with characteristic x_0^{σ} = "Task for decision";
- in place l_2 – β-token with characteristic x_0^{β} = "General restrictions for error, time".

Generalized net is presented by a set of transitions:

$$A = \{Z_1, Z_2, Z_{3,1}, \ldots, Z_{3,6}, Z_{4,1}, \ldots, Z_{4,6}, Z_5, Z_6\},$$

where transitions describe the following processes:

- Z_1 – Process of decision for the type of neural network;
- Z_2 – Choice the different type of neural networks;
- $Z_{3,1}$ – Calculating the output of the neural network 1;

  ...
- $Z_{3,6}$ – Calculating the output of the neural network 6;
- $Z_{4,1}$ – Calculating the backpropagation of the neural network 1;

  ...
- $Z_{4,6}$ – Calculating the backpropagation of the neural network 6;
- Z_5 – Check for existing of a decision;
- Z_6 – Return to the next iteration.

Let the place l_5 have higher priority than places $l_{4,1}, \ldots, l_{4,6}$.

The first transition has the following form:

$$Z_1 = \langle \{l_1, l_2, l_5, l_{11}\}, \{l_3, l_{4,1}, \ldots, l_{4,6}, l_5\}, R_1, \vee(l_1, l_2, l_5, l_{11})\rangle,$$

where:

$$R_1 = \begin{array}{c|ccccc} & l_3 & l_{4,1} & \ldots & l_{4,6} & l_5 \\ \hline l_1 & True & False & \ldots & False & False \\ l_2 & False & False & \ldots & False & True \\ l_5 & False & W_{5,4,1} & \ldots & W_{5,4,6} & True \\ l_{11} & True & False & \ldots & False & False \end{array},$$

and the predicates in the index matrix have the following meaning:

- $W_{5,4,1}$= "The current characteristic of the token in l_3 shows that the neural network 1 will be used".

  …
- $W_{5,4,6}$= "The current characteristic of the token in l_3 shows that the neural network 6 will be used".

The token that enters in place l_3 obtain characteristic

$$x_{cu}^{\theta} = \text{"Task for decision"}.$$

The current characteristic of the token in l_3 shows that the neural network 1 will be used. The tokens that enter places $l_{4,1}, \ldots\, l_{4,6}$, obtain respectively the characteristics:

$$x_{cu}^{\alpha 1} = "E,t", \ldots, x_{cu}^{\alpha 6} = "E,t",$$

where E is the maximal mean square error of the learning and t is the maximal time for the work of the neural network.

The second transition has the form:

$$Z_2 = \langle \{l_3\}, \{l_{6,1},\ldots, l_{6,6}\}, R_2, \vee(l_3)\rangle,$$

where

$$R_2 = \begin{array}{c|ccc} & l_{6,1} & \ldots & l_{6,6} \\ \hline l_3 & W_{3,6,1} & \ldots & W_{3,6,6} \end{array},$$

and

- $W_{3,6,1}$ = "The task is sent for decision in neural network 1"; …
- $W_{3,6,6}$ = "The task is sent for decision in neural network 6";

The tokens that enter places $l_{6,1}$, …, $l_{6,6}$ do not obtain any new characteristics.

$$Z_{3,1} = \langle l_{6,1}, l_{7,2}\}, \{l_{7,1}, l_{7,2}\}, R_{3,1}, \vee(l_{6,1}, l_{7,2})\rangle,$$

where

$$R_{3,1} = \begin{array}{c|cc} & l_{7,1} & l_{7,2} \\ \hline l_{6,1} & False & True \\ l_{7,2} & W_{7,2,7,1} & True \end{array},$$

and $W_{7,2,7,1}$ = "The outputs of the neural network 1 are calculated".

The token that enters place $l_{7,1}$ obtains characteristic "Values of the output of the neural network 1".

The interim transitions from $Z_{3,2}$ $Z_{3,5}$ are formed by analogy. Finally,

$$Z_{3,6} = \langle l_{6,6}, l_{7,12}\}, \{l_{7,11}, l_{7,12}\}, R_{3,6}, \vee(l_{6,6}, l_{7,12})\rangle,$$

where

$$R_{3,6} = \begin{array}{c|cc} & l_{7,11} & l_{7,12} \\ \hline l_{6,6} & False & True \\ l_{7,12} & W_{7,12,7,11} & True \end{array},$$

and $W_{7,12,7,11}$= "The outputs of the neural network 6 are calculated".

The token that enters place $l_{7,11}$ obtains characteristic "values of the output of the neural network 6".

$$Z_{4,1} = \langle l_{7,1}, l_{4,1}, l_{8,2}\}, \{l_{8,1}, l_{8,2}\}, R_{4,1}, \vee(l_{7,1}, \wedge(l_{4,1}, l_{8,2}))\rangle,$$

where

$$R_{4,1} = \begin{array}{c|cc} & l_{8,1} & l_{8,2} \\ \hline l_{7,1} & False & True \\ l_{4,1} & False & True \\ l_{8,2} & W_{8,2,8,1} & W_{8,2,8,2} \end{array},$$

and

- $W_{8,2,8,1}$= "($e_1 < E$ & $T_1 < t$) or $T_1 > t$" ;
- $W_{8,2,8,2}$= $\neg W_{8,2,8,1}$

where

- e_1 – "Mean square error for the current learning of the neural network 1";
- T_1 – "Time for the current learning of the neural network 1".

If $e_1 < E$ & $T_1 < t$, the token that enters place $l_{8,1}$ obtains characteristic “There is a decision of the neural network 1”. If $T_1 > t$, the token that enters place $l_{8,1}$ obtains characteristic “There is no decision of the neural network 1”.

The interim transitions from $Z_{4,2}$ to $Z_{4,5}$ are formed by analogy. Finally,

$$Z_{4,6} = \langle\{l_{7,11}, l_{4,6}, l_{8,12}\}, \{l_{8,11}, l_{8,12}\}, R_{4,6}, \vee(l_{7,11}, \wedge(l_{4,6}, l_{8,12}))\rangle,$$

where

$$R_{4,6} = \begin{array}{c|cc} & l_{8,11} & l_{8,12} \\ \hline l_{7,11} & False & True \\ l_{4,6} & False & True \\ l_{8,12} & W_{8,12,8,11} & W_{8,12,8,12} \end{array},$$

and

- $W_{8,12,8,11}$= “($e_6 < E$ & $T_6 < t$) or $T_6 > t$”;
- $W_{8,12,8,11}$= $\neg W_{8,12,8,12}$

where

- e_6 – “Mean square error for the current learning of the neural network 6”;
- T_6 – “Time for the current learning of the neural network 6.

If $e_6 < E$ & $T_6 < t$, the token that enters place $l_{8,11}$ obtains characteristic “There is a decision of the neural network 6”.

If $T_6 > t$, the token that enters place $l_{8,11}$ obtains characteristic “There is not a decision of the neural network 6”.

The next transition has the form:

$$Z_5 = \langle\{l_{8,1}, l_{8,3}, \ldots, l_{8,11}\}, \{l_9, l_{10}\}, R_5, \vee(l_{8,1}, l_{8,3},\ldots, l_{8,11})\rangle,$$

where

$$R_5 = \begin{array}{c|cc} & l_9 & l_{10} \\ \hline l_{8,1} & W_{8,1,9} & W_{8,1,10} \\ l_{8,3} & W_{8,3,9} & W_{8,3,10} \\ \vdots & \vdots & \vdots \\ l_{8,11} & W_{8,11,9} & W_{8,11,10} \end{array},$$

and

- $W_{8,2i-1,9}$ = “Last characteristic of the current token is that there is a decision of the neural network i, $i \in \{1, 2, \ldots, 6\}$”;
- $W_{8,2i,10} = \neg W_{8,2i-1,9}$

The tokens do not obtain any characteristic in the output places.

The final transition in the generalized net has the form:

$$Z_6 = \langle\{l_9, l_{13}\}, \{l_{11}, l_{12}, l_{13}\}, R_6, \vee(l_9, l_{13})\rangle,$$

where

$$R_6 = \begin{array}{c|ccc} & l_{11} & l_{12} & l_{13} \\ \hline l_9 & False & False & True \\ l_{13} & W_{13,11} & W_{13,12} & False \end{array},$$

and

- $W_{13,11}$ = “There is no token in places $l_{8,1}$,… $l_{8,12}$, and l_9”;
- $W_{13,12} = \neg W_{13,11}$

The token that enters place l_{13} obtains characteristic “the best results among the worst results of the tokens, entering place l_{13} in the current iteration”, while the tokens in places l_{11} and l_{12} do not obtain any new characteristics.

Conclusion

Here we constructed a Generalized Net that represents six neural networks using the ‘backpropagation’ algorithm, which is the classical one in cases of supervised learning. The choice of the networks is made on the basis of the network’s structure.

When determining the degree of training, the time needed for the neural networks’ learning is taken into consideration, together with the mean squared error.

References

[1] Atanasov, K., S. Sotirov, A. Antonov, Generalized net model for parallel optimization of feed-forward neural network, *Advanced Studies in Contemporary Mathematics*, Vol. 15, No. 1 2007, 109–119.

[2] Atanassov, K. *On Generalized Nets Theory*. “Prof. M. Drinov” Academic Publ. House, Sofia, 2007.

[3] Atanassov, K., *Generalized Nets*, World Scientific. Singapore, 1991.

[4] Atanassov, K., On a new hierarchical operator over the Generalized Nets. Part III, Proc. of the Third Int. Workshop on Intuitionistic Fuzzy Sets and Generalized Nets, Warszawa, 7–8 Sept 2003, 35–40.

[5] Atanassov, K., S. Sotirov, Optimization of a neural network of self-organizing maps type with time-limits by a generalized net, *Advanced Studies in Contemporary Mathematics*, Vol. 13, 2006, No. 2, 213–220.

[6] Krawczak, M., S. Sotirov, K. Atanassov, *Multilayer Neural Network Modelling by Generalized Nets*, Warsaw School of Information Technologies, 2010.

[7] Krawczak, M., Generalized Net Models of Multilayer Neural Network, *Advanced Studies in Contemporary Mathematics*, Vol. 7, 2003, No. 1, 69–86.

[8] Krawczak, M., *Multilayer Neural systems and Generalized Net Models*, Akademicka Oficyna Wydawnicza EXIT, Warszawa, 2003.

[9] Krawczak, M., Aladjov, H., Generalized Net Model of Adjoint Neural Network, *Advanced Studies in Contemporary Mathematics*, Vol. 7, 2003, No. 1, 19–32.

[10] Nikolova, M., K. Atanassov, V. Vasilev, Generalized nets as tools for optimization of real processes. *Advanced Studies in Contemporary Mathematics*, Vol. 9, 2004, No. 1, 47–62.

[11] Radeva, V., M. Krawczak, E. Choy, Review and bibliography on generalized nets theory and applications. Advanced Studies in Contemporary Mathematics, Vol. 4, No. 2, 173–199.

[12] Sotirov, S., Generalized net model of the time delay neural network, Issues in *Intuitionistic Fuzzy Sets and Generalized Nets*, Volume 9, Warszawa, 2010, 125–131.

[13] Sotirov, S., M. Krawczak, Modelling Layered Digital Dynamic Network by a Generalized Net, Issues in IFS and GNs, Volume 9, Warszawa, 2011, 84–91.

[14] Sotirov, S., Krawczak M. Modeling the work of self organizing neural network with generalized nets, *Issues in Intuitionistic Fuzzy Sets and Generalized Nets*, Volume 3, Warsaw, 2006, 57–64.

[15] Sotirov, S., Krawczak M., Modeling the algorithm Backpropagation for learning of neural networks with generalized nets. Part 2, *Issues in Intuitionistic Fuzzy Sets and Generalized Nets*, Warsaw, 2003, 65–70.

[16] Sotirov, S., Krawczak M., V. Kodogiannis, Modeling the work of the neural network of type Learning Vector Quantization with generalized net, *Proc. of the 7th Int. Workshop on Generalized Nets*, Sofia, 14-15 July 2006, 39–44.

[17] Sotirov, S., M. Krawczak, V. Kodogiannis. Generalized nets model of the Grossberg neural network. Part 1, Issues in Intuitionistic Fuzzy Sets and Generalized Nets, Vol. 4, 2007, 27–34.