POLSKA AKADEMIA NAUK

Instytut Badań Systemowych

WSPOMAGANIE INFORMATYCZNE ROZWOJU SPOŁECZNO-GOSPODARCZEGO I OCHRONY ŚRODOWISKA

Redakcja:
Jan Studziński
Ludosław Drelichowski
Olgierd Hryniewicz

WSPOMAGANIE INFORMATYCZNE
ROZWOJU
SPOŁECZNO-GOSPODARCZEGO
I OCHRONY ŚRODOWISKA

Polska Akademia Nauk Instytut Badań Systemowych

Seria: BADANIA SYSTEMOWE

tom 36

Redaktor naukowy:

Prof. dr hab. Jakub Gutenbaum

Warszawa 2004

WSPOMAGANIE INFORMATYCZNE ROZWÓJU SPOŁECZNO-GOSPODARCZEGO I OCHRONY ŚRODOWISKA

Redakcja:

Jan Studziński
Ludosław Drelichowski
Olgierd Hryniewicz

Książka wydana dzięki dotacji KOMITETU BADAŃ NAUKOWYCH

Książka zawiera wybór artykułów poświęconych omówieniu aktualnego stanu badań w kraju w zakresie rozwoju modeli, technik i systemów zarządzania oraz ich zastosowań w różnych dziedzinach gospodarki narodowej. Wyodrębnioną grupę stanowią artykuły omawiające aplikacyjne wyniki projektów badawczych i celowych KBN.

Recenzenci artykułów:

Dr Lucyna Bogdan
Prof. dr hab. inż. Olgierd Hryniewicz
Dr Grażyna Petriczek
Prof. dr hab. inż. Andrzej Straszak
Dr inż. Jan Studziński

Komputerowa edycja tekstu: Anna Gostyńska

Wydawca: Instytut Badań Systemowych PAN
ul. Newelska 6, 01-447 Warszawa

Sekcja Informacji Naukowej i Wydawnictw IBS PAN
tel. 836-68-22

Druk: Zakład Poligraficzny Urzędu Statystycznego w Bydgoszczy
Nakład 110 egz.

ISBN 83-85847-92-8
ISSN 0208-8028

GENERALIZED NETS MODELING CONCEPT – NEURAL NETWORKS MODELS

Maciej Krawczak
Systems Research Institute, Polish Academy of Sciences
Warsaw School of Information Technology
<Maciej.Krawczak@ibspan.waw.pl>

In this paper we consider a generalized net interpretation of multilayer neural networks simulation processes. The generalized net methodology developed as a counterpart of Petri nets can be used to model different kind of dynamic systems. In our considered case we developed a generalized net model of a one possible kind of neuron aggregation within each layer. All elements of the complex models are developed.

Keywords: Multilayer Neural Networks, Modelling, Generalized nets.

1. Generalized Nets Models

The basic difference between generalized nets and the ordinary Petri nets is the *place–transition relation* (Atanassov, 1991), in the theory of generalized nets the transitions are objects of a very complex nature. The places are marked by O, and the transitions by ┳.

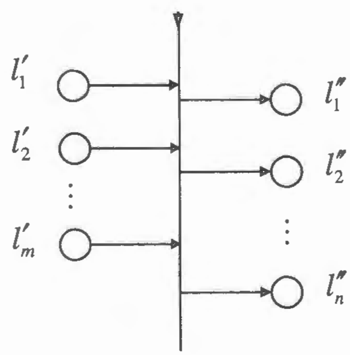

Figure 1: A generalized net transition

Generalized nets contain *tokens*, which are transferred from place to place. Every token bears some information, which is described by token's *characteristic*, and any token enters the net with an *initial characteristic*. After passing a transition the tokens' characteristics are modified.

The transition has *input* and *output* places, as shown in Figure 1.

Formally, every transition is described by a seven-tuple

$$Z = \langle L', L'', t_1, t_2, r, M, \square \rangle \tag{1}$$

where:

- $L' = \{l'_1, l'_2, \dots, l'_m\}$ is a finite, non empty set of the transition's input places,
- $L'' = \{l''_1, l''_2, \dots, l''_m\}$ is a finite, non empty set of the transition's output places,
- t_1 is the current time of the transition's firing,
- t_2 is the current duration of the transition active state,
- r is the transition's *condition* determining which tokens will pass (or will be *transfered*) from the transition's inputs to its outputs; it has the form of an *index matrix* described in (Atanassov, 1987), and briefly recalled in the next section

$$r = \begin{array}{c|ccccc} & l''_1 & \cdots & l''_j & \cdots & l''_n \\ \hline l'_1 & r_{11} & \cdots & r_{1j} & \cdots & r_{1n} \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ l'_m & r_{m1} & \cdots & r_{mj} & \cdots & r_{mn} \end{array}$$

where r_{ij} is a predicate that corresponds to the i-th input and the j-th output places, $1 \le i \le m$, $1 \le j \le n$; when its truth value is *true*, a token is allowed to pass the transition from the i-th input place to the j-th output place,

- M is an index matrix of the capacities of transition's arcs:

$$M = \begin{array}{c|ccccc} & l''_1 & \cdots & l''_j & \cdots & l''_n \\ \hline l'_1 & m_{11} & \cdots & m_{1j} & \cdots & m_{1n} \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ l'_m & m_{m1} & \cdots & m_{mj} & \cdots & m_{mn} \end{array}$$

where $m_{ij} \ge 0$ are natural numbers;

- $\square$ is an object of a form similar to a Boolean expression, it may contain as variables the symbols that serve as labels for transition's input places, and $\square$ is an expression built up from variables and the Boolean connectives $\wedge$ and $\vee$, whose semantics is defined as follows

- $\wedge(l_{i1}, l_{i2}, \dots, l_{iu})$ - every place $l_{i1}, l_{i2}, \dots, l_{iu}$ must contain at least one token,
- $\vee(l_{i1}, l_{i2}, \dots, l_{iu})$ - there must be at least one token in every place $l_{i1}, l_{i2}, \dots, l_{iu}$,

where $\{l_{i1}, l_{i2}, \dots, l_{iu}\} \subset L'$;

when the value of a type (calculated as a Boolean expression) is *true*, the transition can become active, otherwise it cannot.

The following ordered four-tuple

$$E = \left\langle \langle A, \pi_A, \pi_L, c, f, \Theta_1, \Theta_2 \rangle, \langle K, \pi_k, \Theta_K \rangle, \langle T, t^0, t^* \rangle, \langle X, \Phi, b \rangle \right\rangle \qquad (2)$$

is called *generalized net* if the elements are described as follows:

- A is a set of transitions,
- π_A is a function yielding the priorities of the transitions, i.e. $\pi_A : A \to N$, where $N = \{0, 1, 2, \dots\} \cup \{\infty\}$,
- π_L is a function specifying the priorities of the places, i.e. $\pi_L : L \to N$, where $L = pr_1 A \cup pr_2 A$, and $pr_i X$ is the i-th projection of the n-dimensional set, where $n \in N$, $n \geq 1$ and $1 \leq i \leq n$ (obviously, L is the set of all generalized net places),
- c is a function providing the capacities of the places, i.e. $c : L \to N$,
- f is a function that calculates the truth values of the predicates of the transition's conditions (for the generalized net described here let the function f have the value *false* or *true*, i.e. a value from the set $\{0, 1\}$),
- Θ_1 is a function specifying the next time-moment when a given transition Z can be activated, i.e. $\Theta_1(t) = t'$, where $pr_3 Z = t$, $t' \in \lfloor T, T + t^* \rfloor$ and $t \leq t'$; the value of this function is calculated at the moment when the transition terminates its functioning,
- Θ_2 is a function yielding the duration of the active state of a given transition Z, i.e. $\Theta_2(t) = t'$, where $pr_4 Z = t \in \lfloor T, T + t^* \rfloor$ and $t' \geq 0$; the value of this function is calculated at the moment when the transition starts its functioning,
- K is the set of the generalized net's tokens,
- π_K is a function specifying the priorities of the tokens, i.e. $\pi_K : K \to N$,
- Θ_K is a function producing the time-moment when a given token can enter the net, i.e. $\Theta_K(\alpha) = t$, where $\alpha \in K$ and $t \in \lfloor T, T + t^* \rfloor$,

- T is the time-moment when the generalized net starts functioning; this moment is determined with respect to a fixed (global) time-scale,
- t^0 is an elementary time-step, related to the fixed (global) time-scale,
- t^* is the duration of the generalized net functioning,
- X is the set of all initial characteristics the tokens can receive on entering the net,
- Φ is a characteristic function that assigns new characteristics to every token when it makes the transfer from an input to an output place of a given transition,
- b is a function specifying the maximum number of characteristics a given token can receive, i.e. $b: K \to N$.

2. The Concept of Index Matrices

In the description of the transitions in the generalized nets the *index matrices* (Atanassov, 1987, 1997) are used, which will be introduced here in brief.

Let I be a fixed set of indices and R be the set of real numbers. By an index matrix (IM) with index sets K and L $(K, L \subset I)$ we will mean the object:

$$[K, L, \{a_{k_i}, l_j\}] \equiv \begin{array}{c|cccc} & l_1 & l_2 & \cdots & l_n \\ \hline k_1 & a_{k_1,l_1} & a_{k_1,l_2} & \cdots & a_{k_1,l_n} \\ k_2 & a_{k_2,l_1} & a_{k_2,l_2} & \cdots & a_{k_2,l_n} \\ \vdots & & & & \\ k_m & a_{k_m,l_1} & a_{k_m,l_2} & \cdots & a_{k_m,l_n} \end{array}$$

(or briefly: $[K, L, \{a_{k_i,l_j}\}]$) where $K = \{k_1, k_2, \ldots, k_m\}$, $L = \{l_1, l_2, \ldots, l_n\}$, for $1 \le i \le m$, and for $1 \le j \le n$: $a_{k_i,l_j} \in R$.

For the index matrices $A = [K, L, \{a_{k_i,l_j}\}]$, $B = [P, Q, \{b_{p_r,q_s}\}]$ the usual Matrix operations *addition* and *multiplication* are defined, as well as the following operations:

a) $A + B = [K \cup P, L \cup Q, \{c_{t_u,v_w}\}]$.

where

$$c_{t_u,v_w} = \begin{cases} a_{k_i,l_j}, & \text{if } t_u = k_i \in K \text{ and } v_w = l_j \in L - Q \\ & \text{or } t_u = k_i \in K - P \text{ and } v_w = l_j \in L \\ b_{p_r,q_s}, & \text{if } t_u = p_r \in P \text{ and } v_w = q_s \in Q - L \\ & \text{or } t_u = p_r \in P - K \text{ and } v_w = q_s \in Q \\ a_{k_i,l_j} + b_{p_r,q_s}, & \text{if } t_u = k_i = p_r \in K \cap P \\ & \text{and } v_w = l_j = q_s \in L \cap Q \\ 0, & \text{otherwise} \end{cases};$$

b) $A \text{ x } B = [K \cap P, L \cap Q, \{c_{t_u,v_w}\}]$,

where

$$c_{t_u,v_w} = a_{k_i,l_j} \cdot b_{p_r,q_s}, \text{ for } \quad t_u = k_i = p_r \in K \cap P \text{ and } v_w = l_j = q_s \in L \cap Q$$

c) $A.B = [K \cup (P - L), Q \cup (L - P), \{c_{t_u,v_w}\}]$,

where

$$c_{t_u,v_w} = \begin{cases} a_{k_i,l_j}, & \text{if } t_u = k_i \in K \text{ and } v_w = l_j \in L - P \\ b_{p_r,q_s}, & \text{if } t_u = p_r \in P - L \text{ and } v_w = q_s \in Q \\ \sum\limits_{l_j = p_r \in L \cap P} a_{k_i,l_j} b_{p_r,q_s}, & \text{if } t_u = k_i \in K \text{ and } v_w = q_s \in Q \\ 0, & \text{otherwise} \end{cases}$$

d) $A - B = [K - P, \; L - Q, \; \{c_{t_u,v_w}\}]$,

where "–" is the set-theoretic difference operation and

$$c_{t_u,v_w} = a_{k_i,l_j}, \text{ for } t_u = k_i \in K - P \text{ and } v_w = l_j \in L - Q.$$

3. Generalized Nets Models of Neural Networks

A multilayer neural network consists of a number of simple processing units called neurons. The neurons are arranged in L layers, each layer is composed of $N(l)$ neurons, $l = 0,1,2,...,L$, where $N(0)$ denotes the number of inputs. The output of the network is equivalent to all the neurons' outputs from the last L-th layer. The network output is strictly related to the presented input, subject to the conditions resulting from the constancy of the structure (the neuron connections), the activation functions as well as the weights. In this way the neural networks realize the following simulation, that is

$$output = NN(input). \tag{3}$$

The simulation process of neural networks can be modelled by *generalized nets* methodology. In the review and bibliography on generalized nets theory and applications of Radeva, Krawczak and Choy (2002) we can find a list 353 scientific works related to the generalized nets.

The neurons, which consists the neural network can be aggregated in many ways, e.g. any neuron is treated as a subsystem, or the neurons are aggregated within each laser, or neither separate neurons nor layers are distinguished. Here we consider only the first case treating each neuron as a separate subsystem. In this case the considered neural network consists of NL subsystems (neurons) described by the activation function as follows

$$x_{pj(l)} = f\left(net_{pj(l)}\right), \tag{4}$$

where

$$net_{pj(l)} = \sum_{i=1}^{N(l-1)} w_{i(l-1)j(l)}\, x_{pi(l-1)}, \tag{5}$$

while $x_{pi(l-1)}$ denotes the output of the i-th neuron with respect to the pattern p, $p = 1, 2, \ldots, P$, and the weight $w_{i(l-1)j(l)}$ connects the i-th neuron from the $(l-1)$-st layer with the j-th from the l-th layer, $j = 1, 2, \ldots, N(l)$, $l = 1, 2, \ldots, L$.

The generalized net model of the considered aggregation case is shown in Figure 2.

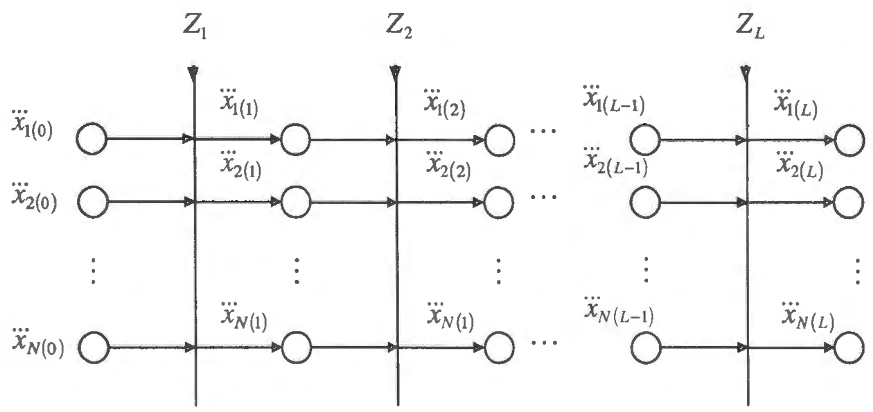

Figure 2. The generalized net model of neural network simulation

The model consists of a set of L transitions, each transition being of the following form

$$Z_l = \left\langle \left\{ \dddot{x}_{1(l-1)}, \dddot{x}_{2(l-1)},\ldots, \dddot{x}_{N(l-1)} \right\}, \left\{ \dddot{x}_{1(l)}, \dddot{x}_{2(l)},\ldots, \dddot{x}_{N(l)} \right\}, \tau_l, \tau'_l, r_l, M_l, \Box_l \right\rangle \quad (6)$$

for $l = 1,2,\ldots,L$, where

$\left\{ \dddot{x}_{1(l-1)}, \dddot{x}_{2(l-1)},\ldots, \dddot{x}_{N(l-1)} \right\}$ - is the set of input places of the l -th transition,

$\left\{ \dddot{x}_{1(l)}, \dddot{x}_{2(l)},\ldots, \dddot{x}_{N(l)} \right\}$ - is the set of output places of the l -th transition,

τ_l – is the time when the l -th transition is fired out, while it is assumed that $\tau_1 = T$ and $\tau_l = T + \sum_{k=2}^{l} \tau'_{k-1}$,

τ'_l – is the duration time of firing of the l -th transition,

r_l – denotes the l -th transition condition determining the transfer of tokens from the transition's inputs $\left\{ \dddot{x}_{1(l-1)}, \dddot{x}_{2(l-1)},\ldots, \dddot{x}_{N(l-1)} \right\}$ to its outputs $\left\{ \dddot{x}_{1(l)}, \dddot{x}_{2(l)},\ldots, \dddot{x}_{N(l)} \right\}$, and has the following index matrix form:

$r_l =$

	$\dddot{x}_{1(l)}$	$\dddot{x}_{2(l)}$	$\ldots$	$\dddot{x}_{N(l)}$
$\dddot{x}_{1(l-1)}$	*true*	*true*	$\ldots$	*true*
$\dddot{x}_{2(l-1)}$	*true*	*true*	$\ldots$	*true*
$\vdots$	$\vdots$	$\ldots$	$\ldots$	$\vdots$
$\dddot{x}_{N(l-1)}$	*true*	*true*	$\ldots$	*true*

(7)

where the value *true* indicates that the tokens representing the neurons can be transferred from the i -th input place to the j -th output place, $i = 1,2,\ldots, N(l-1)$, $j = 1, 2,\ldots, N(l)$,

M_l – indicates an index matrix describing the capacities of transition's arcs:

$M_l =$

	$\dddot{x}_{1(l)}$	$\dddot{x}_{2(l)}$	$\ldots$	$\dddot{x}_{N(l)}$
$\dddot{x}_{1(l-1)}$	1	1	$\ldots$	1
$\dddot{x}_{2(l-1)}$	1	1	$\ldots$	1
$\vdots$	$\vdots$	$\ldots$	$\ldots$	$\vdots$
$\dddot{x}_{N(l-1)}$	1	1	$\ldots$	1

(8)

$\square_l$ – has a form of Boolean expression $\wedge\left(\ddot{x}_{1(l-1)}, \ddot{x}_{2(l-1)},\ldots, \ddot{x}_{N(l-1)}\right)$ and stipulates that each input place $\ddot{x}_{i(l-1)}$, $i=1,2,\ldots,N(l-1)$, must contain a token that will be transferred to the l-th transition.

The generalized net describing the considered neural network simulation process has the following form:

$$GN1=\left\langle \left\langle A, \pi_A, \pi_X, c, g, \Theta_1, \Theta_2\right\rangle, \left\langle K, \pi_k, \Theta_K\right\rangle, \left\langle T, t^0, t^*\right\rangle, \left\langle Y, \Phi, b\right\rangle \right\rangle \quad (9)$$

where

- $A=\{Z_1, Z_2,\ldots, Z_L\}$ - is the set of transitions,
- π_A – is a function classifying the transitions, this classification giving the priorities of the transitions, i.e. $\pi_A : A \rightarrow N$, where $N=\{0, 1, 2,\ldots\} \cup\{\infty\}$ - in the considered neural network case this function is not valid because the transitions are arranged in a natural way (the asterisk * will be used in the subsequent text in order to denote the components of the General net structure which can be omitted),
- π_X – is a function describing the priorities of the places in the following way:

$$pr_1\{Z_1, Z_2,\ldots, Z_L\}= \left\{\ddot{x}_{1(0)}, \ddot{x}_{2(0)},\ldots, \ddot{x}_{N(0)}, \ddot{x}_{1(1)}, \ddot{x}_{2(1)},\ldots, \ddot{x}_{N(1)},\ldots, \ddot{x}_{1(L-1)}, \ddot{x}_{2(L-1)},\ldots, \ddot{x}_{N(L-1)}\right\} \quad (10)$$

$$pr_2\{Z_1, Z_2,\ldots, Z_L\}= \left\{\ddot{x}_{1(2)}, \ddot{x}_{2(2)},\ldots, \ddot{x}_{N(2)}, \ddot{x}_{1(2)}, \ddot{x}_{2(2)},\ldots, \ddot{x}_{N(2)},\ldots, \ddot{x}_{1(L)}, \ddot{x}_{2(L)},\ldots, \ddot{x}_{N(L)}\right\} \quad (11)$$

$$pr_1 A \cup pr_2 A = \left\{\ddot{x}_{1(0)}, \ddot{x}_{2(0)},\ldots, \ddot{x}_{N(0)}, \ddot{x}_{1(1)}, \ddot{x}_{2(1)},\ldots, \ddot{x}_{N(1)}, \ldots, \ddot{x}_{1(2)}, \ddot{x}_{2(2)},\ldots, \ddot{x}_{N(2)},\ldots, \ddot{x}_{1(L)}, \ddot{x}_{2(L)},\ldots, \ddot{x}_{N(L)}\right\}, \quad (12)$$

- c – is a function describing the capacities of the places; in our case it is equal 1 , for $i=1, 2,\ldots, N(l)$, $l=0,1,2,\ldots,L$,
- g – is a function that calculates the truth values of the predicates of the transition conditions, in the considered case $g\left(r_{l,\, i(l-1) j(l)}\right)=true$,
- Θ_1 – is a function yielding the next time-moment when the transitions can be again activated,

- Θ_2 – is a function giving the duration of activity of a given transition Z_l ,
- K – is the set of tokens entering the generalized net, in the considered case there are $N(0)$ input places and each place contains one token; this set can be written as

$$K = \{\alpha_{1(0)}, \alpha_{2(0)}, \ldots, \alpha_{N(0)}\}, \tag{13}$$

- π_K – is a function describing the priorities of the tokens, here all tokens have the same priorities, and it will be denoted by * for $\pi_K(\alpha_{l(0)})$

$l = 1, 2, \ldots, N(0)$,

- Θ_K – is a function giving the time-moment when a given token can enter the net, i.e. all the tokens enter the considered generalized net at the same moment T,
- T – is the time when the generalized net starts functioning – here it is assumed that the net starts at the moment T, when the tokens enter the net,
- t^0 – is an elementary time-step, here this parameter is not used and is denoted by *,
- t^* – determines the duration of the generalized net functioning, that is $t^* = \sum_{l=1}^{L} \tau_l'$,
- Y – denotes the set of all the initial characteristics of the tokens, the characteristics of tokens describe the information which is carried by tokens and changed in transitions,

$$Y = \{y(\alpha_{1(0)}), y(\alpha_{2(0)}), \ldots, y(\alpha_{N(0)})\} \tag{14}$$

where

$$y(\alpha_{i(0)}) = \langle NN1, N(0), N(1), imX_{i(0)}, imW_{i(0)}, F_{(1)}, imout_{i(0)} \rangle \qquad (15$$

is the initial characteristic of the token $\alpha_{i(0)}$ that enters the place $\ddot{x}_{i(0)}$, $i = 1, 2, \ldots, N(0)$, where

$NN1$ – the neural network identifier,

$N(0)$ – the number of input places to the net as well as to the transition Z_1 (equal to the number of inputs to the neural network),

$N(1)$ – the number of the output places of the transition Z_1 ,

$$imX_{i(0)} = [0,\dots,0,x_{i(0)},0,\dots,0]^T \tag{16}$$

- is the index matrix, indicating the inputs to the network, of dimension $N(0)\times 1$ in which all elements are equal 0 except for the element i whose value is equal $x_{i(0)}$ (the i-th input of the neural network),

$$imW_{i(0)} = \begin{array}{c|cccc} & \dddot{x}_{1(1)} & \dddot{x}_{2(1)} & \dots & \dddot{x}_{N(1)} \\ \hline \dddot{x}_{i(0)} & w_{i(0)1(1)} & w_{i(0)2(1)} & \dots & w_{i(0)N(1)} \end{array} \tag{17}$$

- has a form of an index matrix and denotes the weights connecting the i-th input with all neurons allocated to the 1-st layer

$$F_{(1)} = \left[f_{1(1)}\left(\sum_{i=1}^{N(0)} x_{i(0)}\, w_{i(0)1(1)} \right), \dots, f_{N(1)}\left(\sum_{i=1}^{N(0)} x_{i(0)}\, w_{i(0)N(1)} \right) \right]^T \tag{18}$$

- denotes a vector of the activation functions of the neurons associated with the 1-st layer

$$imout_{i(0)} = \begin{array}{c|cccc} & \dddot{x}_{1(1)} & \dddot{x}_{2(1)} & \dots & \dddot{x}_{N(1)} \\ \hline \dddot{x}_{i(0)} & x_{i(0)}\, w_{i(0)1(1)} & x_{i(0)}\, w_{i(0)2(1)} & \dots & x_{i(0)}\, w_{i(0)N(1)} \end{array} \tag{19}$$

- is an index matrix describing the signal outgoing from the i-th input place, $i = 1,2,\dots,N(0)$, to all output places of the Z_1 transition,

- Φ – is a characteristic function that generates the new characteristics of the new tokens after passing the transition; for the transition Z_l, $l = 1,2,\dots,L$, there are $N(l-1)$ input places $\{\dddot{x}_{1(l-1)}, \dddot{x}_{2(l-1)},\dots, \dddot{x}_{N(l-1)}\}$ and with each place there is associated a single token $\alpha_{i(l-1)}$, $i = 1,2,\dots,N(l-1)$, having the characteristic

$$y(\alpha_{i(l-1)}) = \langle NN1,\, N(l-1), N(l),\, imX_{i(l-1)},\, imW_{i(l-1)},\, F_{(l)},\, imout_{i(l-1)} \rangle \tag{20}$$

where

$NN1$ – the neural network identifier,

$N(l-1)$ – the number of input places to the net as well as to the transition Z_l,

$N(l)$ – the number of the output places of the transition,

$$imX_{i(l-1)} = [0,\dots,0,x_{i(l-1)},0,\dots,0]^T \tag{21}$$

- the index matrix of dimension $N(l-1)\times 1$ in which all elements are equal 0 except the element i whose value is equal $x_{i(l-1)}$ - the i-th input value associated with the Z_l transition,

$$imW_{i(l-1)} = \begin{array}{c|cccc} & \dddot{x}_{1(l)} & \dddot{x}_{2(l)} & \dots & \dddot{x}_{N(l)} \\ \hline \dddot{x}_{i(l-1)} & w_{i(l-1)1(l)} & w_{i(l-1)2(l)} & \dots & w_{i(l-1)N(l)} \end{array} \qquad (22)$$

- is an index matrix describing the weight connection between the i-th input places with all output places of the Z_l transition,

$$F_{(l)} = \left[f_{1(l)}\left(\sum_{i=1}^{N(l-1)} x_{i(l)}\, w_{i(l-1)1(l)} \right), \dots, f_{N(l)}\left(\sum_{i=1}^{N(l-1)} x_{i(l-1)}\, w_{i(l-1)N(l)} \right) \right]^T \qquad (23)$$

- is a vector of the activation functions of the neurons associated with the l-th layer of the neural network

$$imout_{i(l-1)} = \begin{array}{c|cccc} & \dddot{x}_{1(l)} & \dddot{x}_{2(l)} & \cdots & \dddot{x}_{N(l)} \\ \hline \dddot{x}_{i(l-1)} & x_{i(l-1)}\, w_{i(l-1)1(l)} & x_{i(l-1)}\, w_{i(l-1)2(l)} & \cdots & x_{i(l-1)}\, w_{i(l-1)N(l)} \end{array} \qquad (24)$$

- is an index matrix describing the signals outgoing from the i-th input place, $i = 1,2,\dots,N(l)$, to all output places of the Z_l transition.

The tokens $\alpha_{i(l-1)}$, $i = 1,2,\dots,N(l-1)$, passing the transition Z_l vanish, and the new tokens $\alpha_{j(l)}$, $j = 1,2,\dots,N(l)$, associated with the output places $\{\dddot{x}_{1(l)}, \dddot{x}_{2(l)}, \dots, \dddot{x}_{N(l)}\}$ of the transition Z_l are generated, their characteristics being described as follows

$$y(\alpha_{j(l)}) = \left\langle NN1, N(l), N(l+1),\ imX_{j(l)},\ imW_{j(l)},\ F_{(l+1)},\ imout_{j(l)} \right\rangle \qquad (25)$$

for $l = 1,2,\dots,L-1$, while

$$y(\alpha_{j(L)}) = \left\langle NN1, N(L), imX_{j(L)} \right\rangle, \qquad (26)$$

and for these new tokens the values $x_{j(l)}$, $j = 1,2,\dots,N(l)$, are calculated in the following way

$$imX_{j(l)} = f_{j(l)}\left(\sum_{i=1}^{N(l-1)} imout_{i(l-1)} \right), \qquad l = 1,2,\dots,L. \qquad (27)$$

It should be mentioned here that $imX_{j(L)}$, $j=1,2,...,N(L)$, denotes the output of the neural network, the final state of the network after ending the simulation process,

- b – is a function describing the maximum number of characteristics a given token can receive; in the here considered neural network simulation process this function has a simple form

$$b(\alpha_{j(l)})=1\,,\ \text{for}\quad j=1,2,...,N(l),\, l=1,2,...,L\,, \tag{28}$$

which means that the characteristic of each token $\alpha_{j(l)}$, $j=1,2,...,N(l)$, $l=1,2,...,L$, is constructed on the base of the characteristics of all tokens ($i=1,2,...,N(l-1)$) from the previous layer $(l-1)$, $l=1,2,...,L$.

Due to the above considerations the transitions have the following form

$$Z_l=\left\langle \left\{\dddot{x}_{1(l-1)}, \dddot{x}_{2(l-1)},..., \dddot{x}_{N(l-1)}\right\}, \left\{\dddot{x}_{1(l)}, \dddot{x}_{2(l)},..., \dddot{x}_{N(l)}\right\},\ \tau_l, \tau_l', *, *, \square_l \right\rangle \tag{29}$$

for $l=1,2,...,L$.

The reduced form of the generalized net describing the simulation process of the neural network has the following form:

$$GN1=\left\langle\, \langle A, *, \pi_X, c, *, \Theta_1, \Theta_2\rangle,\ \ \langle K, *, \Theta_K\rangle,\ \ \left\langle T, *, t^*\right\rangle,\ \ \langle Y, \Phi, b\rangle\,\right\rangle. \tag{30}$$

where

- $A=\{Z_1, Z_2,..., Z_L\}$ - is a set of transitions,
- π_X - is a function describing the priorities of the places,
- c - is a function describing the capacities of the places, i.e. $c(x_{i(l)})=1$, $i=1,2,...,\ N(l)$, $l=0,1,2,...,L$,
- Θ_1 - is a function yielding the next time-moment when the transitions can be again activated, $\Theta_1(t_l)=t_l'$, $l=1,2,...,L$,
- Θ_2 - is a function giving the duration of activity of the transition Z_l $\Theta_2(t_l)=t_l''$, $l=1,2,...,L$,
- $K=\{\alpha_{1(0)}, \alpha_{2(0)},..., \alpha_{N(0)}\}$ - is the set of tokens entering the generalized net,
- $\Theta_K=T$ - for all tokens entering the net and at this moment the net starts to function,
- t^* - determines the duration of the generalized net's functioning and is described by (6.28) or (6.29),

- Y - denotes the set of all initial characteristics of the tokens described by $Y = \{y(\alpha_{1(0)}), y(\alpha_{2(0)}), ..., y(\alpha_{N(0)})\}$, where

  $$y(\alpha_{i(0)}) = \langle NN1, N(0), N(1), x_{i(0)}, W_{i(0)(1)}, F_{(1)}, out_{i(0)} \rangle,$$

- Φ - is a characteristic function that generates the new characteristics of the new tokens after passing the transition,
- $b(\alpha_{j(l)}) = 1$, for $j = 1,2,...,N(l)$, $l = 1,2,...,L$, - is a function describing the number of characteristics memorized by each token.

Such generalized nets with some components missing (the components not being valid) are called *reduced generalized nets* (Atanassov, 1991). In the above version of the generalized nets representation of the simulation process of multilayer neural network we preserve the parallelism of computation.

4. Conclusions

We have described the concept of a generalized nets methodology for modelling discrete event systems, and next the concept of index matrix usefull for aggregation as well as for separation of subsystems. Next, as an example we considered the generalized net model representing the functioning of the multilayer neural networks - the simulation process of this class of networks, and in somehow informal way we have applied many of the sophisticated tools of the generalized nets theory.

References

Atanassov K. (1987) Generalized index matrices. *Competes Rendus de l'Academie Bulgare des Sciences*, **40**, 11, 15-18.

Atanassov K. (1991) *Generalized nets*. World Scientific, Singapore, New Jersey, London.

Atanassov K. (1997) *Generalized Nets and Systems Theory*. „Prof. M. Drinov" Academic Publishing House, Sofia.

Hassoun M. H. (1995) *Fundamentals of Artificial Neural Networks*. MIT Press.

Krawczak M. (1997) Neural networks learning as a multiobjective optimal control problem. *Mathware and Soft Computing*, **4**, 3, 195-202.

Krawczak M. (2000) Backpropagation versus dynamic programming approach. *Bulletin of Polish Academy of Sciences*, **48**, 2, 167-180.

Krawczak M. (2003) *Multilayer Neural Systems and Generalized Net Models*. Academic Press EXIT, Warsaw, Poland.

Krawczak M., Aladjov H. (2002) Generalized Net Model of Backpropagation Algorithm, in: *Proc. of 3rd International Workshop on Generalized Nets*, Sofia, Bulgaria, October 2002, 32-36.

Radeva V., Krawczak M., Choy E. (2002) Review and bibliography on generalized nets theory and applications. *Advanced Studies in Contemporary Mathematics*, **4**, 2, 173-199.

ISSN 0208-8028
ISBN 83-85847-92-8

W celu uzyskania bliższych informacji i zakupu dodatkowych egzemplarzy
prosimy o kontakt z Instytutem Badań Systemowych PAN
ul. Newelska 6, 01-447 Warszawa
tel. 837-35-78 w. 241 e-mail: biblioteka@ibspan.waw.pl