NIFS 4 (1998) 1,15-22

Intuitionistic Fuzzy Generalized Net Model of Machine Learning Process.

Hristo Tsvetaov Aladjov
CLBME-Bulg. Academy of Sci., ad MRL, P.O.Box 12, Sofia-1113, Bulgaria
aladjov@rocketmail.com

Introduction

The aim of the present paper is to construct a Generalized Net (GN-[1,2]) describing the process of machine learning. The task for developing such a model is formulated at the "Session of the Mathematical Foundations of Artificial Intelligence" seminar in 1991 (published in [3]), as a part of the idea that different areas of Artificial Intelligence (AI) should be described with a uniform formalism. The acceptance of such a universal description tool would make it possible to exchange ideas among the different areas, and to concentrate the efforts for solving common problems.

For this purpose the report contains: basic definitions related to the concept of GN; intuitionistc fuzzy GN model for process of machine learning; ideas for development and practicable application.

Generalized Nets

Generalized Nets are an extension of Petri Nets. The modelled object is represented by a combination of states and transitions between them.
Transitions are denoted by Υ and states are represented by places denoted by ○.

The presence of a given object with certain characteristics in a given state is denoted by drawing a point in the corresponding place. On certain conditions transitions become active ('fire') and objects pass from one state to another. The following formal definition of this concept can be given.

Every transition is described by a seven-tuple

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

L' and **L''** are finite, non-empty sets of places called input and output places respectively.

t_1 is the current value of the time - moment of transition firing.

t_2 is the current value of the duration of its active state.

r is the transition condition determining which tokens will pass from transition's input to its output places. It has the form of an Index Matrix (IM) shown on *Fig.1.*

M is an IM of the capacities of transition's arcs, shown on *Fig.2.*

$\square$ is the transition type, determining what combination of tokens the transition's input places should contain so that the transition could become active. Similarly to the Boolean operations "and" and "or" operations $\wedge$ (every place must contain at least one token), $\vee$ (there must be at least one token in all places) are defined.

r	L''_1	...	L''_j	...	L''_n
L'_1					
:					
L'_i			$r_{i,j}$		
:					
L'_m					

Fig.1.

M	L''_1	...	L''_j	...	L''_n
L'_1					
:					
L'_i			$M_{i,j}$		
:					
L'_m					

Fig.2.

The ordered four-tuple

$$\mathbf{E=<<A, \pi_A, \pi_L, c, f, \theta_1, \theta_2>,<K, \pi_K, \theta_K>,<T,t^{\circ},t^*>,<X, \phi, b>>}$$

is called a Generalized Net if:

A is a set of transitions.
π_A is a function that gives the transitions' priorities π_A : E -> *N*.
π_L is a function giving the priorities of the places π_L : L -> *N*.
c is a function giving the capacities of the places .
f is a function which calculates the truth values of the predicates .
θ_1 is a function giving the next time moment when a given transition become active. $\theta_1(t) = t`$, where $t, t` \in [T,T+t^*], t < t`$.
θ_2 is a function giving the duration of the active state of a given transition $\theta_2(t) = t`$, where $t, t` \in [T,T+t^*], t` > 0$.
K is the set of the GN's tokens.
π_K is a function giving the priorities of the tokens π_K: K -> *N*.
θ_K is a function giving the time-moment when a given token can enter the net. $\theta_K(\alpha) = t$, where $\alpha \in \hat{E}, t \in [T,T+t^*]$.
T is the time moment when the GN starts functioning.
t° is an elementary time step.
t^* is the duration of the GN's functioning.
X is the set of the initial characteristics the tokens can receive when they enter the net.
ϕ is a characteristic function which assigns new characteristics to every token when it makes transfer from an input to an output place of a given transition.
b is a function giving the maximum number of characteristics a given token can receive
b: K -> *N*. Where $N=\{0,1,....\} \cup \infty$

In the described model we use Intuitionistic fuzzy GNs of type 1 [1,2], in which predicates are evaluated into set $[0,1] \times [0,1]$. Or in different words the function **f** is defined as follows:

$$f(r_{i,j}) = \langle \mu(r_{i,j}), \nu(r_{i,j}) \rangle, \text{ where}$$

$\mu(r_{i,j})$ is the truth degree of the predicate $r_{i,j}$, and $\nu(r_{i,j})$ is its falsity degree and

$$\mu(r_{i,j}) + \nu(r_{i,j}) \le 1$$

Standard notations are used for predicates that have zero fuzziness:

$$f(\text{true}) = f(T) = f(1) = \langle 1,0 \rangle$$
$$f(\text{false}) = f(F) = f(0) = \langle 0,1 \rangle$$

Generalized Net Model Describing Process of Machine Learning

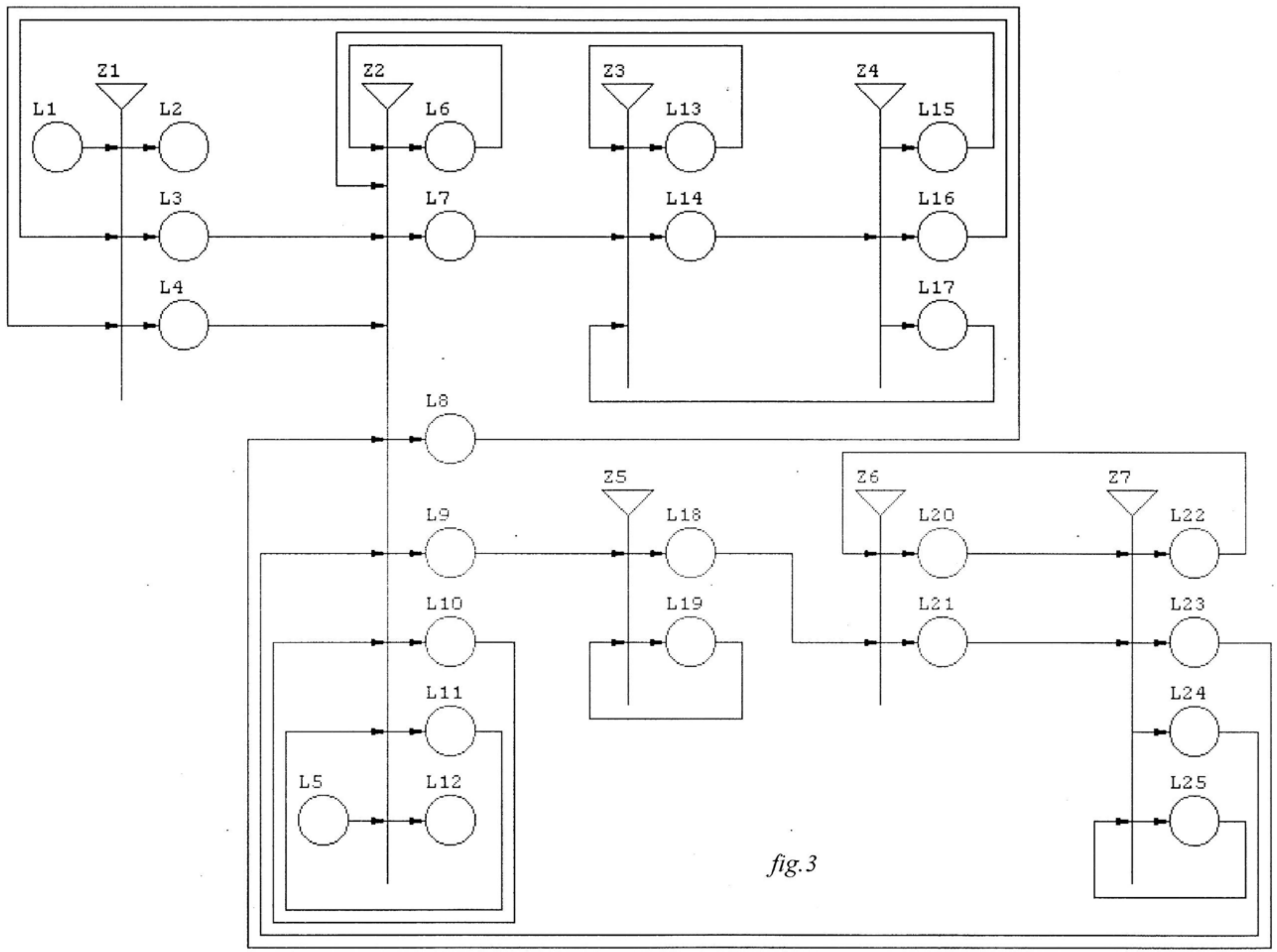

fig.3

Intuitionistc Fuzzy Generalized Net Model for Process of Machine Learning

GN model is shown in *Fig.3*. It consists of objects and procedures for training that are represented by α and β tokens. Each α token has its Intuitionistic fuzzy characteristic, $\langle \mu, \nu \rangle$, where μ and ν mean the degree of training and degree of untraining, respectively. Every β token has such characteristic for each object trained by a procedure which represents that token. This characteristic determines procedure training efficiency.

The learning process is divided into steps represented by transitions Z1, Z2 ... Z7. The α tokens enter model via L1 place. After that they are evaluated with respect to their training. For trained tokens are considered, such tokens whose degree of training μ is above a given threshold, and their untraining degree is under a given threshold. Trained objects leave the net via L2, and the rest of them continue their training by entering in L3 or L4. The value of predicate W3, where they will get into, and if they get into L4 then the training procedures are changed. The necessity of such change depends on how the object's training degree and untraining degree are changed since the last iteration of the given procedure. The formal definition of Z1 is given below.

$$Z1 = \langle \{L1, L8, L16\}, \{L2, L3, L4\}, r_{Z1}, M_{Z1}, \vee(L1, L8, L16) \rangle, \text{ where}$$

r_{Z1}	L2	L3	L4
L1	W_1	$\overline{W_1} \cdot W_2$ [1]	$\overline{W_1} \cdot W_3$
L8	$\overline{W_3}$ [2]	false	W_3
L16	W_1	$\overline{W_1} \cdot W_2$	$\overline{W_1} \cdot W_3$

M_{Z1}	L2	L3	L4
L1	$M_{1,2}$	$M_{1,3}$	$M_{1,4}$
L8	$M_{8,2}$	$M_{8,3}$	$M_{8,4}$
L16	$M_{16,2}$	$M_{16,3}$	$M_{16,4}$

and:

W_1 ="the object is enough trained"

W_2 ="the object continues its training"

W_3 ="training procedures must be adjusted"

The rest of the transitions' parameters $\pi_L(L1)$, $\pi_L(L16)$, f, c, θ_1, θ_2, π_K, θ_K represent learning process features that can be varied to fit the limitations of real implementations or in the process of searching optimal training scheme with respect to efficiency. For example it depends on $\pi_L(L1)$ and $\pi_L(L16)$ whether the new or already trained tasks have priority, or function "c" can be given the capacity of in-buffer. With tokens characteristic π_K can be assigned priorities to training objects. The kind of θ_1, θ_2 will be changed depending on whether the estimated criteria can be used in parallel or in sequence.

---

[1] where " $\cdot$ " is logical "and"

[2] $\overline{W_3}$ means negation of W_3

From L3 token can enter L6 if there are more than one parallel executable group of procedures. In this passage (transition) it obtain as a characteristic all applicable, parallel executable procedures. After that when token passes form L6 to L7, in its characteristic there remain only groups of compatible procedures $\phi = \phi(X_0^\alpha, X_0^\beta, t, \phi(t-1))$. In case that the subset of assigned procedures is executable, the tokens are split. First group of tokens begin training by free subset of group of procedures and they enter L14. Second group of tokens wait looping in L13 until procedures are released. Which token will continue its training can be determined by a great number of disciplines, discipline kind can also be object of simulation. For example tokens priorities can be used, or the duration of their stay in queue, or whether they are already trained by another procedures, or whether objects instance are trained in the moment.

In Z4 transition tokens obtain new intuitionistic fuzzy rating for their training. Corresponding to that rating can be determined whether the object continues its training by its group of procedures, or it must change its procedures, or it must come back to Z1 where must be evaluated its closeness to wished training stage. Simultaneously with that is evaluated procedure training efficiency by tracing change of object's training before and after a training-step. In procedures contour this evaluation is made by changing efficiency characteristics corresponding to given object to active procedures that loop in L10. Criteria by which the influence of each procedure is determined individually and their work together over the object is of intersection. Intuitionistic fuzzy interpretation of procedure's efficiency allow cases of impossible determination which procedure brings given change to be represented with the indefiniteness of a intuitionistic fuzzy characteristic. The possibility for parallelly working procedures must be envisaged when their parameters are modified, because they can be optimized with respect of their parallel work or for their independent work separately.

In parallel with described contour function and contour of training procedures represented with β tokens. They enter model via L5 place and in case that are needless they leave model via L12. Procedures that are currently used for training loop in L10. Inactive procedures loop in L11. If the procedures characteristics should be modified, tokens enter L9. In case when there is more than one parameter by which the procedure will be evaluated, the token (that represents this procedure) splits into two tokens and the first of them enters L19, the second enters L18. Thus L18 contains tokens with characteristic concrete evaluation parameter, which can be for example the achieved velocity of training, the number of simultaneously trained objects with a given efficiency

Every procedure has an intuitioniste fuzzy efficiency rating with respect to each trained object. On this basis the overall procedure efficiency can be evaluated. In Z6 procedures are considered efficient which enter L20, or as inefficient which enter L21. For achieving better efficiency procedures are specialized to train a subset of objects from currently trained ones. I. e. procedures are tuned to trained objects with similar fuzzy characteristics. This process is represented by splitting the tokens passed from L21 (L25) to L25. The opposite process of generalizing procedures optimized for training certain group of tokens is accomplished by merging the procedures heaped in L20 into one that enters L23. It must be taken into account that in this process procedures may have various raitings by different criteria.

$$Z2 = \left\langle \begin{matrix} \{L3, L4, L5, L6, L10, L11, L15, L23, L24\}, \{L6, L7, L8, L9, L10, L11, L12\}, \\ r_{Z1}, M_{Z1}, \vee(L3, L4, L5, L6, L10, L11, L15, L23, L24) \end{matrix} \right\rangle,$$

where

r_{Z2}	L6	L7	L8	L9	L10	L11	L12
L3	W_4	$\overline{W_4}$	false	false	false	false	false
L4	false	false	true	false	false	false	false
L6	false	true	false	false	false	false	false
L5	false	false	false	$W_3 \cdot \overline{W_6}$	$\overline{W_6} \cdot W_7$	$\overline{W_6} \cdot \overline{W_7}$	W_6
L10	false	false	false	$W_3 \cdot \overline{W_6}$	$\overline{W_6} \cdot W_7$	$\overline{W_6} \cdot \overline{W_7}$	W_6
L11	false	false	false	$W_3 \cdot \overline{W_6}$	$\overline{W_6} \cdot W_7$	$\overline{W_6} \cdot \overline{W_7}$	W_6
L15	W_4	$\overline{W_4}$	false	false	false	false	false
L23	false	false	false	$W_3 \cdot \overline{W_6}$	$\overline{W_6} \cdot W_7$	$\overline{W_6} \cdot \overline{W_7}$	W_6
L24	false	false	false	$W_3 \cdot \overline{W_6}$	$\overline{W_6} \cdot W_7$	$\overline{W_6} \cdot \overline{W_7}$	W_6

and

W_4 ="there are parallel executable group of procedures"

W_6 ="useless procedures"

W_7 ="used for training"

This model does not require specific values of arcs capacities and thier values in transitions' definintions will be omitted.

$$Z3 = \langle \{L7, L13, L17\}, \{L13, L14\}, r_{Z3}, M_{Z3}, \vee(L7, L13, L14) \rangle, \text{ where}$$

r_{Z3}	L13	L14
L7	$\overline{W_5}$	W_5
L13	$\overline{W_5}$	W_5
L17	$\overline{W_5}$	W_5

and

W_5 ="there is free subset of group of procedures from assigned ones"

$$Z4 = \langle \{L14\}, \{L15, L16, L17\}, r_{Z4}, M_{Z4}, L14 \rangle, \text{ where}$$

r_{Z4}	L15	L16	L17
L14	W_{12}	W_{13}	W_{14}

and

W_{12} ="training procedures must be changed"

W_{13} ="it needs evaluation of training stage" or "it needs modification of training procedure parameteres "

W_{14} ="it needs a subsequent training step"

$$Z5 = \langle \{L9, L19\}\{L18, L19\}, r_{Z5}, M_{Z5}, \vee(L9, L19) \rangle, \text{ where}$$

r_{Z5}	L18	L19
L9	$\overline{W_{11}}$	W_{11}
L19	$\overline{W_{11}}$	W_{11}

and

W_{11} ="there are more parameters for valuation"

$$Z6 = \langle \{L18, L22\}, \{L20, L21\}, r_{Z6}, M_{Z6}, \vee(L18, L22) \rangle, \text{ where}$$

r_{Z6}	L20	L21
L18	W_8	$\overline{W_8}$
L22	W_8	$\overline{W_8}$

where W_8 means

W_8 ="the procedure is efficient by the given criteria"

$$Z7 = \langle \{L20, L21, L25\}, \{L22, L23, L24, L25\}, r_{Z7}, M_{Z7}, \vee(L20, L21, L25) \rangle, \text{ where}$$

r_{Z7}	L22	L23	L24	L25
L20	W_9	true	false	false
L21	W_9	true	W_{10}	W_{10}
L25	W_9	true	W_{10}	W_{10}

and

W_9 = "procedures nust be modified"

W_{10} ="procedures must be specialised"

Conclusion

The exhibited model can be considered as one of the possible models of machine learning process. It allows one to monitor the work of training procedures, study their efficiency at different stages of learning process, and to examine their mutual influence in their parallel work. Many detailisations can be done by using hierarchical operators that replace a given transition or place with a whole sub-net with the same behavior. Most of the models parameters can be considered as tokens characteristics from an additional contour that models the training process course. In this contour optimizations can be performed with respect to some given purpose or stastistic information can be collected.

Bibliography

[1] Atanassov, K. Generalized Nets in Artificial Inteligence Vol. 1 Generalized Nets and Expert Systems, Sofia, Publishing House of the Bulgarian Academy of Science "Prof. Marin Drinov", 1998.

[2] Atanassov, K. Introduction to the Theory of Generalized Nets. Pontica-Print, Bourgas, 1992. (in Bulgarian)

[3] Attanassov, K. Generalized nets and artificial intelligence.- Advances in Modelling & Analysis, B, Vol. 37, 1997, No 1-2,37-51