\

NEURAL MODELS OF FUZZY SET CONNECTIVES

[1]Kaoru Hirota [2]Witold Pedrycz
[1]Dept. of Instrument and Control Eng.
Hosei University, Koganei-city, Tokyo 184 Japan

[2]Dept.of Electrical and Computer Eng.
Univ.of Manitoba, Winnipeg, Canada R3T 2N2
pedrycz@eeserv.ee.umanitoba.ca

Abstract The paper introduces a neural network based model of logical connectives. The network consists of two types of generic OR and AND neurons structured into a three layer topology. The specificity of the logical connectives is captured by the network within its supervised learning. Further analysis of the connections of the network obtained in this way provides a better insight into the nature of the connectives for fuzzy sets; in particular the analysis can well focus on their non-monotonic and compensative properties. Numerical studies carried out for the Zimmermann-Zysno data set illustrate the performance of the network.

Keywords logical connectives (operators), logic-based neurons, OR/AND neuron, learning, nonmonotonic reasoning

1. Introduction

The question of modelling generic logical operations (connectives) on fuzzy sets has attracted attention from the very early stage of development of the area. Both the advancements of the theoretical foundations and experimental verification of diverse models have been pursued. One may refer to an axiomatic approach put forward in [1], a variety of models examining the use of triangular norms [2] as being two representative streams of the theoretical investigations. The experimental results are scarce. Nevertheless, the experiments reported in [10] clearly highlighted a higher semantic complexity of the fuzzy set operations that it had been anticipated. In particular, it has been revealed that the "pure" AND or OR character of the operations as conveyed by most of the existing set-theoretic models available nowadays cannot cope well with the available experimental data. Admitting that, in [10] the authors alleviated the problem by introducing a new compensatory logic operator whose characteristics constitute a mixture of AND and OR features while the contribution of those two can be conveniently modelled by an auxiliary parameter of the model. The model of this connective takes on the form

$$y=[AND(x_i,x_j)]^{1-\gamma}\,[OR(x_i,x_j)]^{\gamma}$$

where x_i and x_j stand for the grades of membership being aggregated, y denotes the result of this aggregation, while γ is a weight factor used to express a grade of compensation, $\gamma\in[0,1]$. The value of γ equal to 0 yields no compensation while γ set to 1/2 produces the highest compensation between the AND and OR types of the aggregation. The AND and OR operations could be realized as the standard maximum and minimum functions.

Another version of the connective exhibiting a similar compensatory nature can be completed as the following convex combination,

$$y=\gamma AND(x_i,x_j)+(1-\gamma)OR(x_i,x_j)$$

If we accept the minimum and bounded sum as the corresponding AND and OR operations, the above expression is read as

$$y=\gamma\min(x_i,x_j)+(1-\gamma)\min(1,x_i+x_j)$$

The logic operator of this form has been investigated in [6].

The above two-argument operators can be directly expanded into their many-variable versions.

The weighted mean has been studied as another example of the compensative logic operator [3]. For n-arguments it reads as

$$y=\left(\sum_{i=1}^{n} c_i x_i^p\right)^{1/p}$$

$p\in\mathbf{R}$, $p\neq 0$. Here the modifiable degree of compensation is accomplished by changing the power "p". Depending on its value the model embraces several situations: for $p=-\infty$ it yields the minimum operation while for $p=+\infty$ it behaves as the maximum operator. One can also choose a monotonic mapping between the corresponding values of p and γ in the previous model.

The data reported in [10] constitute now a benchmark that is used to experiment with new models of the logical operations. For instance, the compensatory operations have been recently studied in the context of neural networks applied towards their realization, cf. [5].

In [9] a new model of nonmonotonic logic operations has been introduced. This study calls for a substantial revision of the existing models in order to make them capable of handling the principles of nonmonotonic reasoning.

In this paper we will propose a new model of logical connectives that in its realization uses exclusively standard set-theoretic operations (triangular norms) whereas a compensative character can be achieved by developing some structural relationships between them. From a computational point of view, the model can be also treated as a certain class of logic-based neuron that is constructed with the aid of AND and OR logic neurons. In general, the proposed model is nonmonotonic. The issue of "local " domain-dependent behaviour of the logical connectives, as emphasized in [1], will be addressed through the parametric learning of the connections of the network. Furthermore, this learning makes it possible to quantify numerically the effect of non-monotonicity of the operator.

First we will study the model itself along with its features and learning. This discussion is followed by the detailed numerical studies.

2. OR/AND neuron as a model of logical connectives

Before proceeding with the detailed architecture and learning realized for the overall network, we will briefly remind the two basic types of logic-based neurons as proposed in [7] [8].

The AND neuron aggregates input signals (membership values) $\mathbf{x} = [x_1\ x_2 \ldots x_n]$ by first combining them individually with the connections (weights) $\mathbf{w}=[w_1\ w_2 \ldots w_n]$ and afterwards globally ANDing these partial results,

$$y=\mathrm{AND}(\mathbf{x};\mathbf{w}),$$

i.e.,

$$y = \mathop{T}_{i=1}^{n} [x_i\ t\ w_i]$$

where t- and s-norms are used to represent AND and OR operation, respectively.

The structure of the OR neuron is dual to that reported for the AND neuron, namely,

$$y=\mathrm{OR}(\mathbf{x};\mathbf{w}),$$

that reads coordinatewise as

$$y = \overset{n}{\underset{i=1}{S}} [x_i \, t \, w_i]$$

The AND and OR neurons realize "pure" logic operations on the membership grades. The role of the connections is to distinguish between different level of impact the individual inputs might have on the result of aggregation. Considering the boundary conditions of the triangular norms we conclude that higher values of the connections in the OR neuron emphasize the stronger influence that the corresponding inputs have on the output. The opposite effect takes place in the case of the AND neuron: the values of w_i close to 1 make an influence of x_i almost negligible.

Analysing a list of general postulates formulated for the models of fuzzy set connectives such as continuity, monotonicity, associativity, and commutativity, we can learn that the two first from this list are automatically preserved as a straightforward consequence of the utilization of the triangular norms. The commutativity (and subsequently associativity) requirement is, in general, not satisfied since the connections of the neurons establish a certain priority between the inputs. The forfeiture of commutativity is characteristic for logical operations used in nonmonotonic reasoning, cf. [4]. The models of non-monotonic connectives as proposed in [9] show clearly this lack of commutativity. For instance, the nonmonotonic AND operator provided there is defined by augmenting the usual monotonic operator x_iANDx_j by an additional nonsymmetrical component $\lambda(x_i, x_j)$ expressing a priority occurring between the arguments. This essentially re-translates into the following form of the expression for the connective,

$$y=\max(\lambda(x_i,x_j), x_i \text{ AND } x_j)$$

To maintain monotonicity one has to keep all the connections equal, $w=w_i$, $i=1,2,\ldots,n$.

The connections modify also the boundary conditions of the logic operations so that they may not coincide with their Boolean counterparts. The admissible range of output values y for the AND neuron is computed by taking all x_i equal to 0 or 1. In virtue of the monotonicity property of the triangular norms, we obtain,

$$y \in [\overset{n}{\underset{i=1}{T}} w_i, 1]$$

Analogously, for the OR neuron one derives the condition,

$$y \in [0, \overset{n}{\underset{i=1}{S}} w_i]$$

Thus the connections different from 0 and 1 limit the lower and upper bound of the values produced by the logical connective. For the AND neuron one can interpret the lower limit as an initial confidence one puts in a simultaneous satisfaction of x_i,s. Note also that this overall nonzero initial confidence requires that it holds for all the arguments (all w_i=0).

In the case of the OR neuron, the upper bound emerges as a direct consequence of our belief that the complete satisfaction of all x_i,s could lead to the highest level of this aggregation y less than 1 (conservative approach). By accepting that position we are tempted to consider a limited credibility that is reflected numerically by the values of the connections lower than 1. It is sufficient, though, that at least one of these connections (levels of belief) is equal to 1 and the result of the aggregation expand into the entire unit interval.

It has been shown in [9] that the monotonicity property might not be preserved in nonmonotonic reasoning. The discussed neurons can well handle this aspect of nonmonotonicity by adding complements of x_i,s, $1-x_i$, as auxiliary inputs of the neurons.

The proposed OR/AND neuron constitutes in fact a three-layer network and is constructed by arranging the discussed neurons into a structure displayed in Fig.1.

The relevant formulas are given as

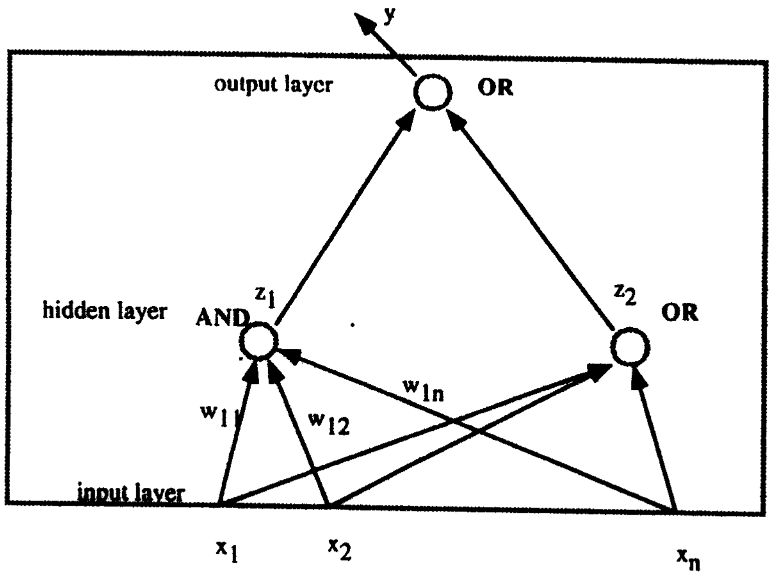


Fig.1. Architecture of OR/AND neuron

$$y=OR([z_1\ z_2];\mathbf{v})$$
$$z_1=AND(\mathbf{x};\mathbf{w}_1) \quad \text{and} \quad z_2=OR(\mathbf{x};\mathbf{w}_2) \tag{1}$$

with $\mathbf{v}=[v_1\ v_2]$, $\mathbf{w}_i=[w_{i1}\ w_{i2}\ldots w_{in}]$, i=1,2.
We can encapsulate the above expressions into a single formula writing down

$$y=OR/AND(\mathbf{x};\mathbf{connections})$$

with **connections** summarizing symbolically all the connections of the network.

The pure characteristics produced at the level of the hidden layer are then combined by the OR neuron constituting the output layer. Note also that by changing the connections between the hidden and the output layer one can easily have an access to the entire range of intermediate characteristics varying from the AND to the OR-like behaviour. The connections **v** provide a necessary flexibility in achieving various levels of compensation between the AND and OR character of the neuron. In particular, the condition $v_1=1$, $v_2=0$ gives rise to a pure AND character of the network (no contribution from the OR neuron). The opposite holds for $v_1=0$ and $v_2=1$.

The issue of learning in the above OR/AND neuron will be addressed as a problem of supervised learning. The updates of the connections $\mathbf{w}=[w_{ij}]$ and $\mathbf{v}=[v_1\ v_2]$ are controlled by the gradient of the considered performance index. Below we will derive detailed formulas considering the Mean Squared Error (MSE) performance criterion and admitting an on-line type of learning (that means that the relevant updates of the connections are performed after a presentation of each pair of the elements of the training data set). Let $(\mathbf{x},t)$ denote a given pair of data from this set. We derive,

$$\Delta w_{ij}=-\alpha/2\frac{\partial Q}{\partial w_{ij}}$$

and

$$\Delta v_i = -\alpha/2 \frac{\partial Q}{\partial v_i}$$

$i=1,2$, $j=1,2,\ldots,n$ where $Q=(t-y)^2$.
Subsequently,

$$\frac{\partial Q}{\partial w_{ij}} = -(t-y)\frac{\partial y}{\partial z_i}\frac{\partial y}{\partial v_i}$$
$$\frac{\partial Q}{\partial v_i} = -(t-y)\frac{\partial y}{\partial v_i}$$

Further detailed computations of the above derivatives can be worked out upon specification of the triangular norms being used to develop the neurons. The enhancements of the learning procedure as explained in [7] can be found applicable here.

3. Numerical experiments

In our experiments we will utilize a benchmark data set coming from [10]. This data set consists of triples (x_1,x_2,t) where x_1 and x_2 are membership values of the two arguments to be logically connected while t denotes the result of this aggregation. For simulation purposes the discussed model (1), n=2, will be further specialized by studying the product and probabilistic sum as examples of t- and s-norms, respectively. Obviously, one can select any other combination of the triangular norms; moreover they need not to be dual. The learning formulas can be written explicitly as,

$$\frac{\partial Q}{\partial w_{1j}} = -(t-y)v_1(1-z_2v_2)A(1-x_j)$$

$$\frac{\partial Q}{\partial w_{2j}} = -(t-y)v_2(1-z_1v_1)(1-B)x_j$$

$$\frac{\partial Q}{\partial v_1} = -(t-y)z_1(1-z_2v_2)$$

$$\frac{\partial Q}{\partial v_2} = -(t-y)z_2(1-z_1v_1)$$

where

$$A=\prod_{l\neq j} (w_{1l}+x_l-w_{1l}x_l)$$

and

$$B=\underset{l\neq j}{S} (w_{2l}x_l)$$

The initial values of the connections have been selected randomly (taken as random numbers drawn according to a uniform probability distribution function defined over [0,1]). The learning rate α was set to 0.1. The successive values of the cumulative performance index (taken as a sum of Q's for the individual input-output pairs) are visualized in Fig.2. The obtained results plotted versus the target values are given in Fig.3.Overall, they fit well the experimental results. The connections of the neurons are equal to
-hidden -to - output layer

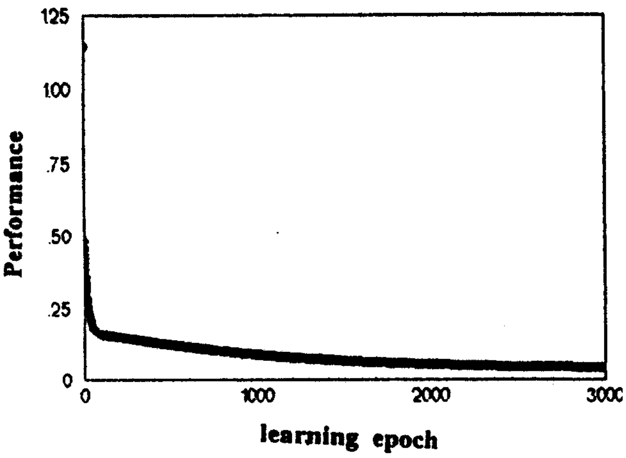
$$v = [\ 0.9957 \quad 0.999\ ]$$

-input -to - hidden layer
AND neuron

$$w_1 = [\ 0.387 \ \ 0.004\ ]$$

OR neuron

$$w_2 = [\ 0.281 \ \ 0.026\ ]$$



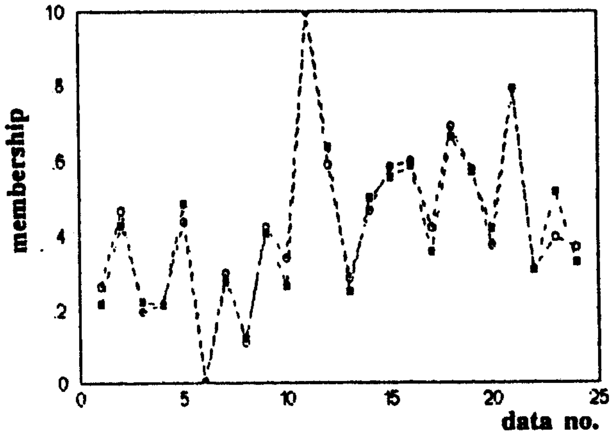


Fig.2. Performance index as a function of successive steps of learning *Fig.3. Membership values in the training set, - ■ data , O-model*

The analysis of the connections reveals that both the neurons in the hidden layer contribute to the same extent to the output. The main difference lies in a substantially asymmetric treatment of the input arguments by the OR and AND neuron. In fact, for the AND neuron the priority assigned to x_2 is higher than that considered for x_2. The order of the priorities for the OR neuron is reversed.

In the next model the commutativity of the logical connective has been retained by keeping the same values of all the connections for the OR and AND neurons of the hidden layer. The constraints imposed in this way produced a slightly higher value of the cumulative performance index (that stabilized at 0.0585 after 3000 learning epochs).The computed connections indicate now a visible distinction between the contributions of the neurons in the hidden layer to the output,
-hidden -to - output layer

$$v = [0.999 \ \ 0.49]$$

-input -to - hidden layer
AND neuron

$$w_1 = [\ 0.02 \ \ 0.02\ ]$$

OR neuron

$$w_2 = [0.616 \ \ 0.616\ ]$$

The results produced by some other models existing in the literature are compiled in the table below

reference	[10]	[6]	[3]	[5]	proposed model
Q (cumulative)	0.080	0.219	0.080 and 0.066	0.044	0.0428 and 0.0585

Considering the accuracy of the measurements of the membership functions, we can conclude that most of the models perform equally well and the produced differences are quite negligible.

4. Conclusions

We have proposed a neutral network (OR/AND neuron) as a new distributed model of logical connectives. Essentially, this network produces nonmonotonic and compensative logic operators. The level of nonmonotonicity could be also controlled at the learning phase. The parametric flexibility of the network will be particularly useful in a detailed quantification of the two features stated above in a given applicational domain and being manifested by the relevant training set. It can be found useful in handling a broad class of problems including those emerging in the area of decision-making.

5. References

1. R.Bellman, L.A.Zadeh, Local and fuzzy logics, In: J.M.Dunn, E.Epstein, eds., Modern Uses of Multiple Valued Logic, Reidel, Dordrecht, 1977, pp. 103-165.
2. H.Dubois, H.Prade, A review of fuzzy set aggregation connectives , Information Sciences,36, 1985, 85-121.
3. H.Dyckhoff, W.Pedrycz, Generalized means as a model of compensative connectives,Fuzzy Sets and Systems, 14, 1984, 143-154.
4. M.L.Ginsberg, Nonmonotonic Reasoning, Morgan Kaufmann, Los Altos, 1987.
5. J.M.Keller, Z.Chen, Learning in fuzzy neural networks, Proc. 2nd Int. Conf. on Fuzzy Logic and Neural Network, Iizuka, Japan, July 17-22,1992, vol.I, pp.85-87.
6. M.K.Luhandjula, Compensatory operators in fuzzy linear programming with multiple objectives, Fuzzy Sets and Systems, 2, 1982, 245-252.
7. W.Pedrycz,"Neurocomputations in relational systems", IEEE Trans. on Pattern Analysis and Machine Intelligence, 13, 1991, pp. 289-296.
8. W.Pedrycz, Fuzzy neural networks and neurocomputations, Fuzzy Sets and Systems, to appear.
9. R.R.Yager, Non-monotonic set theoretic operations, Fuzzy Sets and Systems, 42, 1991, 173- 190.
10. H.J.Zimmermann, P.Zysno, Latent connectives in human decision making, Fuzzy Sets and Systems, 4, 1980, 37-51.