Numerical-Logical Processing in Neural Networks for the Decision Support

Viorel Ariton, Doinita Ariton, “Danubius” University
Galati blvd. 3, Galati 800654, Romania
variton@univ-danubius.ro dariton@univ-danubius.ro

Abstract: Artificial Neural Networks (ANN) embed shallow knowledge through learning. Used in diagnosis and decision support, ANN are immediate computational models for effects and causes as from human experience but keep out from the deep knowledge of them. The paper presents a way of embedding logical processing over the numerical ones in ”neural logical sites” for the classical ANN paradigms, then proposes a way of structuring deep knowledge in the network for all types of abduction problems in a unified way, which is compared with similar way of structuring deep knowledge in the network for all types of neural logical sites for the classical ANN paradigms, then proposes a way of embedding logical processing over the numerical ones. The approach may be spread in any diagnosis and decision support applications involving deep and shallow knowledge.

I. INTRODUCTION

Artificial Neural Network (ANN) paradigm is generally accepted as a modelling attempt to deal with discrete knowledge pieces and the connectionism. Any concepts imagined by people share some characteristics when describing a phenomenon: knowledge is conceived as discrete pieces, and the correlations between them refer to causal links in the given context of the phenomenon, combined in logical relations and put together to reflect aspects of the context for the phenomenon. The discrete pieces of knowledge could be features of the objects (colour, shape, dimension) or “features” of actions (duration, intensity, etc.).

While the various interconnections from concept-neurons reflect shallow knowledge – as the basic neural network model; logical relations between concepts reflect some deep knowledge of the human expert and it is not common to embed them in the neural model, while changing its structure thoroughly.

However, the logical meaning is of crucial importance in the description of real facts, at least for the sake of the “explanation” – so, causal relations are get deterministic models (formula) or they represent simply links between features and/or objects in the frame of a human experts’ explanation.

On the other side, in just the way they where conceived and work, ANN are best fitted for correlations made between various patterns of discrete information. However, due to their intrinsic connectionist nature, they hardly deal with logical processing.

The paper proposes a simple way of embedding logical processing in classical ANN structures, in order to perform some discrimination of contexts when pattern recognition is to be used in decision making. One of the immediate application of the novel ANN structure is used – in the last part of the paper, to emulate abduction problems useful in diagnosis, while multiple relations between effects and causes get straight connectionist models, and the logical relations between causes and between effects to causes, get specific (neural site) implementations for the input functions of the neurons embedded in the ANN structure. Using deep knowledge of human expert about specific abduction problem solving one can configure ANN with “logical neural sites” for any diagnosis expert system using softcomputing techniques.

II. NEURAL MODELS WITH LOGICAL OVERLOAD

Artificial Neural Networks, usually, use simple processing elements interconnected in an exhaustive way, while structured or not on layers. The “artificial” neuron try to act in a similar way as the natural one, i.e. remaining inactive as far as the cumulated stimuli do not surpass a certain threshold, and getting active (fired) after that.

A. The site as the input function of the neuron

In the well known mathematical model of the neuron as in [5], the inputs are weighted and arithmetically summed by the so called input function; after that, the activation function is applied (slope or sigmoid like function), to obtain a pseudo-binary value.

To get closer to the proposed approach below, let consider the inputs and the output of the neuron as concepts involved in a decision process like diagnosis: the plausibility of a cause $C_i$ (the output) increases when the contribution of the effects $E_j$ (the inputs) lead to a reasonable value:

$$C_i = f(\sum_{j=1}^{E_i} w_{ij} \cdot E_j + \theta).$$

in other words, the cause neuron $C_i$ is activated by all the effects $E_j$ from the entire set $E$ (with $|E|$ the set cardinality).

Of course, the effects and the causes may have strong correlations or just zero-correlation among them. The cause neuron $C_i$ is fired according to the activation function $f$ and to the input function of the neuron - the argument of $f$. The cumulative action of all effects $E_j$ indicate a general plausibility of the cause, while each effect evokes the cause in a specific measure indicated by its weight $w_{ij}$. After that, the value of the cumulated effects may fire the neuron if the threshold $\theta$ is surpassed.

The input function is the sum of products in the argument of the function $f$ in (1). Some works consider the input function as...
a separated processing taking place in the so called “site” (see [6]). Primarily, the site processing is the cumulative operation between inputs, and the output of the site is simply the summed value of all inputs. This kind of processing may be considered as a logical OR, while in a logical OR gate it doesn’t matter which input’s contribution is the one leading to the activation of the gate or, in other words, it doesn’t matter which input is more intense.

It is obvious that considering the plausibility of a cause as the contribution of all effects is not quite close to an explanation but similar to the shallow knowledge of a diagnostician – coming merely from practice.

B. Logical overload of the neural sites

Let’s consider that the activation values of \( C_i \) and \( E_j \) falls in \([0, 1]\) interval so, they may get logical (qualitative) meanings as follows: if \( E_j \) equals 0 then the effect is absent, if 1 – it is certain, while 0.5 is the doubt level; the same meanings one can use for a cause \( C_i \).

The activation of an effect \( E_j \) gets a logical meaning as follows: when in \([0, 0.5]\) interval it is “not important” but when in the \((0.5, 1]\) interval it is “important”. When the current activation of the effect is processed by the site of the neuron (the input function) the effect activation is weighted through the link – indicating the correlations of the effect to the cause so, the doubt value is weighed, i.e. input \( I \) obeys the rule:

\[
\text{if } I > w/2 \text{ then } I = \text{“important” } \text{ else } I = \text{“not important”}. \tag{2}
\]

The meaning above will support the logical overload so, it can be handled in logical aggregations, and the site becomes a “logical neural site”. Note that the input function in the logical neural site yet cumulates the inputs in the same manner as in the original input function – so, no numerical change will occur in the functioning of the neuron. The logical meaning is only applied (overloaded) to help some logical meaning useful when deep knowledge on effects and causes can bring a better performance to the plausibility of a cause.

C. Logical aggregation though logical neural sites

As shown above, the input function (hence the site) of the neuron originally looks like a logical OR, while each input \( E_j \) contributes (in a weighted degree) to the neuron’s activation - see the argument of \( f \) in (1). How should it look a logical AND or a logical NOT in a similar approach?

The logical overload proposed above may help in this respect, while the input level is associated to logical FALSE or logical TRUE when it is “not important” or “important” – see (2). Using the logical overload, the sites may proceed to the logical processing (over the native arithmetical ones) as in truth tables in Fig. 1. – which shows the “logical neural sites” AND and NOT.

Introducing a generic neural logical site, one can consider logical aggregations OR, AND and NOT as follows (examples are given for sites defined as in Fig. 1).

![Logical site diagram](image)

**Disjunctive aggregation**, performed by the “disjunctive site” through a default cumulative processing of a neural site, i.e. inputs simply cumulate their activation \( I_j \) ("classical" site):

\[
O = \sum_{j=1}^{2} I_j \tag{3}
\]

**Conjunctive aggregation**, performed by the “conjunction site”, which output \( O \) obeys the truth table from Fig. 1 a, i.e. following the rule:

\[
\text{if } I_1 > w_1/2 \text{ AND } I_2 > w_2/2 \text{ then } O = I_1 + I_2 \text{ else } O = 0 \tag{4}
\]

**Negation**, performed by the “negation site”. The output \( O \) is obtained from the input \( I \) according to Eq. 5, and the truth table in Figure 1b:

\[
O = w - I \tag{5}
\]

The logical sites can now embed deep knowledge of the human diagnostician on effects and causes, if he or she is acquainted with some special logical relations between them; it is obvious that the deep knowledge will lead to a better decision support compared to the “blind” combination of effects to causes in the “classical” neuron / site.

III. NEURAL MODELS OF FOR THE ABDUCTION PROBLEMS

As an application of the proposed “neural logical sites”, the paper shows how straightforward and sound solutions will get the abduction problems – as they were defined first by Bylander in reference [3].

The main reason ANN are used for the abduction problem solving (like diagnosis) is that the multiple interactions between effects and causes (in this order) are best caught by the connectionist models; however, no diagnosis relies only on the links of the effects-to-causes, while some causes may
appear in pears (when one occurs it is compulsory another does) or there are other logical relations between effects and causes also between causes.

In the sequel of the paper, the five abduction problems are presented, along with novel neural models for abduction problems solving using logical neural sites.

A. Five abduction problems

Bylander et al. [3] reveal four categories of abduction problems in diagnosis:
- independent abduction problems – when no interaction exists between causes;
- monotonic abduction problems – an effect appears at cumulative causes;
- incompatibility abduction problems – pair of causes are mutually exclusive;
- cancellation abduction problems – pair of causes cancel some effects, otherwise explained separately by one of them.

Ayeb et al. [2] have a sound approach in the neural network modelling of the abduction problems, also they introduce a fifth category:
- open abduction problems - when observations consist of three sets: present, absent and unknown observations.

The following approach is based on [1], where the abduction consists in sequentially applying plausibility criteria – to obtain the set of causes possibly evoked by the set of present effects, then relevance criteria – to obtain the minimum cardinality subset of causes (the “parsimonious principle” as in [4]) but also other restrictions coming from the running context (e.g. frequency, reliability). In the ANN model, the plausibility criteria are the “excitatory” links from effects to causes, and the relevance criteria are competitions – each induce each by a given restriction.

B. Abduction problem solving by neural models using neural logical sites

The present paper only deals with the plausibility criteria, proposing neural network models for each abduction problem above, in a unified and simple approach.

The ANN model for an abduction problem consists in a structure of neural sites performing the logical aggregation, specific for effects and causes in concern (as \( E \) and \( C \) neurons in Fig. 2). Each structure is placed in the target ANN architecture according to the deep knowledge of the human diagnostician on effects and causes.

Each type of abduction problem is solved in Fig. 2. through a specific structure of neural sites, involving forward links from effects to causes and from causes to effects:
- For independent abduction problems – excitatory links apply directly from the effect \( E_j \) to the corresponding cause \( C_i \) (see Fig. 2 a). If there exists a conjunction grouping of effects to the cause, a conjunction site is provided at the input of the cause neuron. Note that by default, the neuron implements a disjunctive grouping of inputs (sum - Eq. 3), represented by the simple triangle.

![Fig. 2. Each abduction problem is solved by a specific neural structure of sites with logical overload.](image)

- For monotonic abduction problems – the causes \( C_i \) and \( C_l \) both evoke the same effect \( E_j \), hence they suffer conjunction with one-another and with the common effect through conjunction sites, as shown in Fig. 2 b, and expressed by the rule:

\[
C_i \leftarrow C_i \ AND \ E_j, \quad C_l \leftarrow C_l \ AND \ E_j
\]  

(6)

- For incompatibility abduction problems – the pair \( C_i \) and \( C_l \) of causes are mutually exclusive (i.e. they are not both active in the same time), both evoking the same effect \( E_j \). Each of them suffers conjunction with the negation of the other cause and with the common effect, as shown in Fig. 2 c, and expressed by the rule:

\[
C_i \leftarrow NOT \ C_l \ AND \ E_j, \quad C_l \leftarrow NOT \ C_i \ AND \ E_j
\]  

(7)

- For cancellation abduction problems – the pair of causes \( C_i \) and \( C_l \) reduce the effect \( E_j \), when both occurred, although each of them evokes it separately. They suffer conjunctions as in Fig. 2 d, according to the following rule:

\[
C_i \leftarrow C_i \ AND \ NOT \ E_j, \quad C_l \leftarrow C_l \ AND \ NOT \ E_j
\]  

(8)

- For open abduction problems – the main task is dealing with absent effects, so the cause \( C_i \) is activated if no effect \( E_j \) exists (Fig. 2 e), according to:

\[
C_i \leftarrow NOT \ E_j
\]  

(9)

Links between cause-neurons in abduction problems of type b, c, d above, have all weights between cause neurons equal to 1 if
they are symmetric (one to another), else they are set according to deep knowledge of the human expert.

The approach presented above is much more simple and sound, comparing with the “unified models for abduction” proposed by [2], who creates special, complicated, excitatory and inhibitory links between neurons associated to causes and effects to catch all aspects of the abduction problems, and also changes the structure of ANN to meet specific needs for each abduction problem (even adjusting weights during the recall phase!). The proposed approach is much simpler and straightforward, reflecting a real unified and manner for all abduction problems – while it uses a unique approach and sound specific structures for them.

C. Adding the Neural Models for Abduction to an ANN Paradigm

The neural models above may be used in any ANN paradigm in the diagnosis task, when direct links exist between effect-neurons and cause-neurons. Plausibility criteria refer to interactions between effects and causes also between causes, which are embedded in the neural network structures of such ANN, as follows:

- into weights of the forward links between effects evoked and causes – as the shallow knowledge obtained through ANN training on known pairs cause-effects;
- into neural sites structures attached to cause neurons, according to respective abduction problem – as deep the knowledge coming from human experts on specific effect-to-cause and cause-to-cause interactions;
- into threshold of the site – as deep knowledge from human experts (usually set to 0).

When building a neural network meant for diagnosis by abduction, one should be acquainted with the set of effects and causes – regarding their interactions as indicated by the deep knowledge. The chosen neural network paradigm is one with two layers – e.g. Adaline, which exhibits direct links between input and output neurons.

During the training phase, the gate functioning of the sites is disabled, i.e. the “logical overload” is not present and only the “classical” cumulative input function is running. The training procedure runs as usual, adapting weights of the links between effect and cause neurons that correspond to the logical relations between them so, the embedding of “shallow knowledge” takes place. In other words, all the structures involved in the “deep knowledge processing” in not present at the training phase; only after that, the deep knowledge is applied (i.e. the structures of logical sites are added) just to improve the decision process using ANN.

In the recall phase, the logical overload in sites will better reproduce the situation from the training phase (regarding the abduction problems) because, besides shallow knowledge, now the deep knowledge provided in the neural models of the abduction problems will contribute to the cause (output) neurons’ activation.

IV. CASE STUDY ON THE DIAGNOSIS FOR A HYDRAULIC INSTALLATION

The neural models above were added to the ANN Adaline and it was compared to [2] approach – the only one referring to all abduction problems. The approaches were used for the fault diagnosis in a simple hydraulic installation (see Fig. 3). First, some practical considerations then an experiment will be presented.

The installation in Fig. 3. comprises the Supply Unit (consisting in pump, tank and pressure valve), the Hydraulic Brake (control valve, brake’s cylinder), and the Conveyor (control valve, self, the conveyor’s cylinder). It present faults: 2 at the tank, 4 at the pump, 3 at the pressure valve, 2 at the pipes, 2×2 at the control valves, 2 at the damper (Drossel), 2×2 at the hydraulic cylinders.

The installation in Fig. 3. comprises the Supply Unit (consisting in pump, tank and pressure valve), the Hydraulic Brake (control valve, brake’s cylinder), and the Conveyor (control valve, self, the conveyor’s cylinder). It present faults: 2 at the tank, 4 at the pump, 3 at the pressure valve, 2 at the pipes, 2×2 at the control valves, 2 at the damper (Drossel), 2×2 at the hydraulic cylinders.

The neural networks used for the experiment comprised the same number of neurons on the output layer for the 21 faults, and different numbers of neurons on the input layer: 48 observations which are also manifestations for the [2] approach, while in the present approach for a manifestations correspond 2 or 3 neurons, depending on the logical meaning attached: on/off or low/normal/high - respectively.

In the case of independent abduction problems, [2] introduces excitatory links from effects to causes and additional inhibitory links (competition) between causes sharing same effects, eventually freezing weights’ values. In the present approach, competition is a relevance criterion (among others – beside the minimum cardinality). So, it offers a more flexible way on selecting the diagnostic. The minimum cardinality in [2] is always 1, while in the present approach multiple fault diagnosis is allowed.

For monotonic abduction problems, [2] introduces a third layer which combines the incompatibility abduction problems – because of the cause conjunction in the two cases. Here, some links skip the third layer (see the independent problems...
above), some enter the layer; so, the building of the network structure is non-homogenous and difficult. Moreover, the compromise between inhibitory and excitatory links for causes in conjunction, may lead to instability during the training (see Fig. 4 - left); Adaline ANN rapidly converges when effect-neurons attack specific cause-neurons.

Fig. 4. Training of the fault ‘Pump supply pipe clogged’, for [2] (left) and present approaches.

The training procedure is different for different abduction problems in the [2] approach, and it involves also competition (between weights and cause neurons). In the present approach, the original ANN paradigm’s training is kept for all neurons.

In the recall phase the same patterns of effects are applied to the inputs of the two neural networks for the 21 faults. In the case of the 4th fault (‘Pump supply pipe clogged’) the output of the two networks are depicted in Fig. 5.

In both approaches, the deep knowledge on various cause-effect and cause-cause interactions should be obtained from the human expert in the target domain. The deep knowledge embedding is simpler in the present approach, and the neural network structure can be automatically generated, using the building blocks in Fig. 2. In the [2] approach (left), the activation of candidate faults and competition between them take place the same time, while in the present approach the ‘plausibility criteria’ (i.e. neural models for abduction) activate faults as in Fig. 5 right, then the ‘relevance criteria’ (i.e. multiple competitions) will later assert plausible and relevant causes (the 4th fault wins against 5th in Fig. 5).

The case study referred to a simulated behaviour of the target installation and included all types of abduction problems presented above. On the whole, all 21 faults get recognized after the recall phase – consisting in plausibility and relevance.

The reason for the good results is the way plausible faults get obtained, i.e. using the deep knowledge embedded in the proposed neural models for abduction problems. There are still faults, not included in the set of causes that may occur in the target hydraulic installation so, the open space of causes may induce errors in the diagnostic of the real installation.

V. CONCLUSION

The ANN paradigm is primarily a structure that embeds shallow knowledge in a computational model and so, it can put together various concepts which correlations hardly get have an explanation. In the human way of thinking this is a common but for practical reasons deep knowledge is often used to get a better performance to the problem solving.

The paper deals with a novel way of embedding logical meanings and logical operations in the artificial neuron processing (which natively is numerical) so, mixing shallow and deep knowledge as usually human experts use in the
decision process. The simple and straightforward “logical neural site” is a generic gate-like processing means that do not change the original way of the ANN paradigm, while it is applied only after all the learning procedure is finished and the decision has to be supported in the recall phase just to obtain an improved answer to the problem.

The gate-like sites (representing the input functions of the neurons) get specific combinations which are illustrated for the case of all five (known) abduction problems in the literature. So, logical relations between concepts, included into the original ANN, is actually the deep knowledge that human expert may add to his or her experience (shallow knowledge) which was embedded in the ANN during the training phase. Such combination of knowledge types will lead to a better performance in diagnosis or in decision making processes, in fact leading to better performance to any abduction problem solving appearing in the real life.

The connectionist solution for all types of abduction problems solving is presented in the paper, along the comparison to similar approaches in the field of abductive reasoning and the advantages of the current approach.

REFERENCES