Decision Validation and Emotional Layers on Fuzzy Boolean Networks•

José A. B. Tomé
Instituto Superior Técnico/ INESC-ID
Rua Alves Redol nº9, 1000 Lisboa, Portugal
jose.tome@inesc-id.pt

João P. Carvalho
Instituto Superior Técnico/ INESC-ID
Rua Alves Redol nº9, 1000 Lisboa, Portugal
joao.carvalho@inesc-id.pt

***Abstract*. Fuzzy Boolean Networks are capable of learning and reasoning. However, the reasoning result must be validated by an emotional layer, which ensures the acting rules are meaningful and that no contradictory rules are giving a wrong, "averaged", defuzzified result. Here the problem is addressed and an emotional layer is developed to deal with it.**

I. INTRODUCTION

Fuzzy Boolean Nets can learn from real experiments and they are capable to perform qualitative reasoning based on their internal memories, which are set during the learning phase [8]. These nets can be considered a neural fuzzy model where the fuzziness is an inherent emerging property in contrast with other known models where fuzziness is artificially introduced on neural nets or where neural components are inserted on the fuzzy systems [2, 3, 4, 5, 7]. Moreover, it has been proved [9] that they are Universal Approximators (they are equivalent to Parzen Windows estimators [6]) and that they are capable to distinguish between different rules, provided a relationship between the number of rules and the number of inputs per neuron is obeyed [10].

Based on this theoretical background some applications have been implemented [12], but there are interesting questions yet to be developed, such as the one here treated.

Since the learnt rules are distributed among a network of neurons, two important questions arise:

1. The first one is: how can the network, itself, know that it has learnt enough from the real world, in order to trust its output on a future application experiment (or, in other words, is the performed reasoning trustable, based on the teaching)?
2. The second question is related with situations where different activated rules provide different outputs from the same input data of a given experiment, in a dangerously way. This situation tends to set the system output on a kind of "average" of the different rule decisions. This problem arises also on the traditional fuzzy systems, where the process of deffuzification gets the "responsibility" of solving it. It is the known case of the automobile control when this "sees" one obstacle ahead, just in front. If some of the rules drive the automobile to the left, in order to avoid the eminent disaster, and some other rules drive it to the right the result may be catastrophic!

On Fuzzy Boolean Nets the network is by itself and must decide what to do, without any external, artificial or algorithmic help. It is the purpose of this paper to present a natural (embedded on the network) solution to these problems.

The first problem is easily solved if one uses, in addition to the internal activation ratios of the neurons, the non-taught ratio. That is, the output of the system (which is represented by the ratio of "ones" among every output neuron) is provided together with a "credibility" output (also a ratio of activated neurons), which gives an indication about the trust one can put on the output.

The second problem is more interesting and is dealt with another layer of neurons, which inputs do not come from the external variables but from the internal states of the reasoning layer of neurons. This is possible because when reasoning takes place, and due to the internal structure of the binary neurons, the network disposes of all necessary information about how every rule is firing, about the output activation ratio and thus about any incoherence.

That is, one uses a meta-network that interprets the work of the network itself and validates its decisions. Reasoning may be performed but decisions not taken at all, if the meta-network is not satisfied with the process. In a certain way this can be viewed as a neural network that provides "emotion" to the reasoning process itself. It is known that on animals, humans particularly, real decisions can not be achieved without the work of the emotional layers [1], even if reasoning has been performed.

II. EMBEDDING LEARNING KNOWLEDGE

The values taken by the linguistic terms on a given experiment are expressed by the binary memories activation ratios and these determine the consequent area activation ratio (which represents the consequent deffuzification). Then, any initial activation ratios (or corresponding to antecedent rule areas not taught enough) have no meaning but wrongly affect the consequent. In order to deal with this situation it is necessary to memorize, for each joint antecedent area of each neuron (corresponding to a given rule antecedent), not only the Boolean value ("0" if that joint antecedent has not been addressed or "1" if it has) but also extra binary information specifying if any thing at all has or not been taught. In hardware terms this would mean two bits of memory per

• This work is partially supported by FCT project POSI/SRI/33741/2000

neuron and per joint antecedent area, instead of one. Let us designate these two bits by *value* and *credit*, respectively.

With this, the learnt knowledge of each rule, after the teaching phase, should now be interpreted not only by the *active consequent activation ratio* of that rule/joint antecedent area (defined as the ratio between the number of "ones" among every value bit of that rule on every consequent neuron and the number of "ones" among the respective credit bits), but also by the *teaching ratio* of the same rule (defined as the ratio of credit bits of that rule on every consequent neuron and the total number of consequent neurons, since each consequent neuron has one and only one credit bit per rule). Thus after the teaching phase one knows which rules have been taught or not, just by evaluating the teaching ratios of those rules. On the reasoning phase (when antecedent values are applied and one wants to get the consequent value) a similar attitude can be taken. Since each neuron output is just the "ORING" of every value bit of that neuron, *if it has been addressed by the particular joint antecedent sampled configuration*, and since just one value bit is addressed per neuron, the credit bit corresponding to that value bit can also be furnished as an extra output. Thus one gets two output areas – the *active output area* and the *credit output area* – and the output activation ratio of the first may be interpreted with the credit given by the activation ratio of the second. For example, if the net has not been taught at all one could obtain any active output activation ratio (say 80%) without any meaning, because the credit activation ratio was 0%. This can be a valuable improvement to validate or not the reasoning output or to take consciousness of which rules have been or not taught.

On fig.1 one may have an idea of a possible neuron with n antecedents each one with 3 inputs to the neuron. Four different internal rule memories (each one corresponding to a different joint antecedent linguistic term) and each one with 2 bits – value and credit. In such a case (if each antecedent provides 3 inputs to the neuron) one should get 3^n sets of two bit internal memories. It is supposed that the addressed internal rule is Rj (upper right), meaning that the neuron output is "0" and it has been taught (credit "1"). To obtain the output activation ratio one should look to every neuron output. In the same way one could obtain each rule consequent value and the respective credit looking to the internal memories of that rule throughout every neuron.

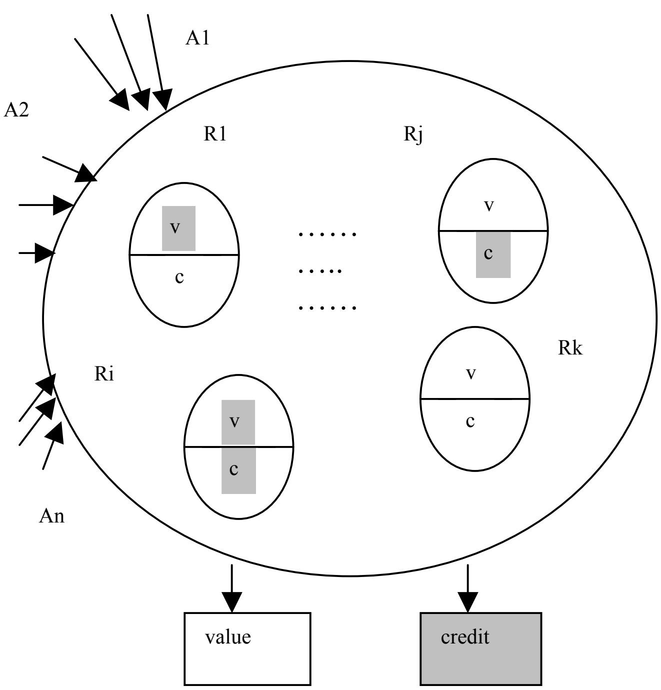

Fig. 1 Internal structure of a Neuron-Value and Credit bits

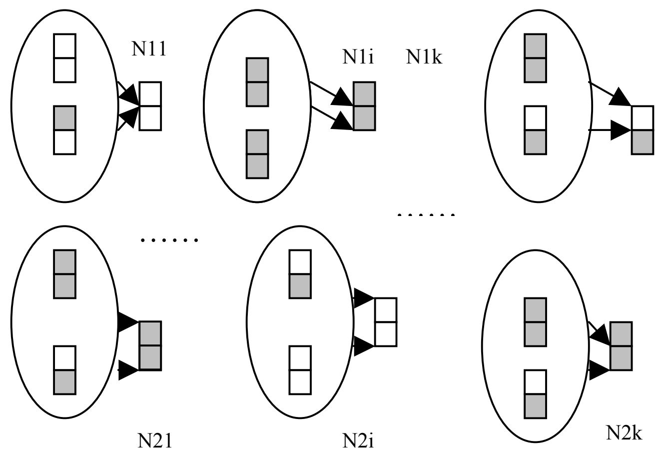

Fig. 2 Example

For example, in fig.2, only two rules (rule A-Above and rule B-Below) are shown in six neurons, as well as their active and credit outputs.

It is supposed that Rule A has been addressed on neurons N11, N1i, N21 and N2k, and Rule B on neurons N1k and N2i. If the activation ratios were to be maintained globally among every neuron, this would mean (since only 5 neurons have credit "1" for that rule and, of those one has the value at "0") that the consequent of Rule A was learnt with a value of 4/5 – probably a linguistic term of High or Very High, depending on the number of membership functions used, and Rule B 1/4. The result of the network (the defuzzication of the results obtained by the 2 rules), on the shown experiment, would be of 3/4 with a credibility of 6/8.

To address the second problem, that is antagonic results from different rules with the same input data, an emotional layer (based on relational operators already developed for these nets [11]) is developed.

III. THE EMOTIONAL LAYER

From the point one has arrived at in the last section, it is known that inside each neuron one has two bits per antecedent rule and that they are established by the teaching process: one, Vi, that gives the "value" of the consequent on what that neuron is concerned, and another, Ci, that assures, if at "1", that the neuron has indeed learnt that rule. In addition, when a real experiment of reasoning happens, each neuron will "see" a joint sample of the inputs coming from the antecedent areas and one and only one of the bit rules of the neuron will be addressed; let one name Di $\in \{0,1\}$, with i being the rule index, the binary variable which is "1" for the addressed rule and "0" for the other rules.

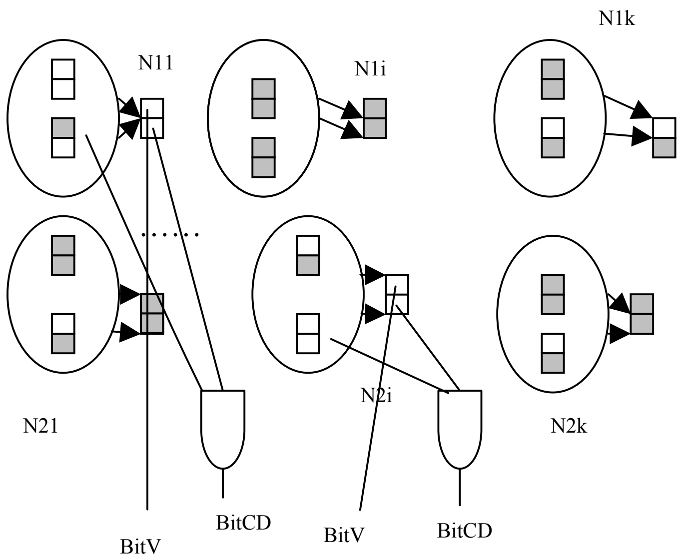

Fig. 3 Bits for Rule B area

Now, if, during the teaching phase, similar input antecedent values have provided very different outputs different consequent have been learnt by the network for the same antecedents. If the actual reasoning experiment is not far (in linguistic terms of the antecedents) from those, this situation tends to set the system output on a kind of "average" of the different rule decisions. This problem arises also on the traditional fuzzy systems, where the process of defuzzification gets the "responsibility" of solving it. It is the known case of the automobile control when this "sees" one obstacle ahead, just in front. If some of the rules drive the automobile to the left, in order to avoid the eminent disaster, and some other rules drive it to the right the result may be catastrophic.

A possible solution for this problem is the use of another neural layer that monitors the internal activity of the other neurons, through the values Vi, Ci and Di. Its responsibility is to determine if two or more rules are been "reasonably" activated and if the outputs of those rules are "quite" different and validated by credit bits. The final decision of the network depends on this layer, and it will not be taken unless permission is granted by it. Another possibility is this "monitoring" layer to decide for one of the two contradictory outputs, based on the relative credits, for example. The kind of behavior of that layer, recalls for what the emotional layers are supposed to do [1], and thus one names it as the emotional layer.

Although the layer can implement more sophisticated differentiations, since one wants to separate rules with "quite" different activations it is enough to use the "low" and "high" activations.

The input "areas" for this layer are composed with the bits of the same rule from every neuron of the reasoning layer. For instance, in fig. 2, the lower bits V_B and C_B together with D_B (the last one not shown but implicit because the output bits of each neuron are those addressed by the antecedent rule present) from neuron N11 to neuron N2k are used to form input area 2.

The logic equation corresponding to the conditions for forming one input (from one consequent neuron) of the emotional input area, corresponding to any rule i, is then:

Bit_i=AND(Di,Ci,Vi), and since the two output bits are given by Vo=$\sum_i$AND(Di,Vi) and Co=$\sum_i$AND(Di,Ci), it comes Bit_i=AND(Vo,Co,Di) because only one Di is "1". This information can be provided in a more useful way if bit Vi is separated from the AND of Di with Ci as it is shown in fig. 3, with two pairs of these bits (the same should be repeated for every neuron), from rule B input area to the emotional layer (in fig. 3 they are denoted by BitV and BitCD, respectively). The set of these bits, for an area/rule i are named as areas Vi and CDi.

The "emotional" layer can be constructed with a generalized relational operator made of the same kind of neurons [11], or it can be decomposed on a series of more simple relational operators, each of them just comparing a pair of rules. In the first case the inputs for that layer come from every pair of areas CDi and Vi. Taking the second possibility, the inputs for each of the sub-networks/relational operators can be the same, but using only the CD and V of the two involved rules. The logical operation such a net must implement is:

IF (AND (Vi "quite different" from Vj, CDi is "activated a lot", CDj is "activated a lot")) Then Block

This kind of reasoning ensures that blocking the output of the reasoning only occurs if both rules have been taught and are being activated on the present experiment (CDi AND CDj)

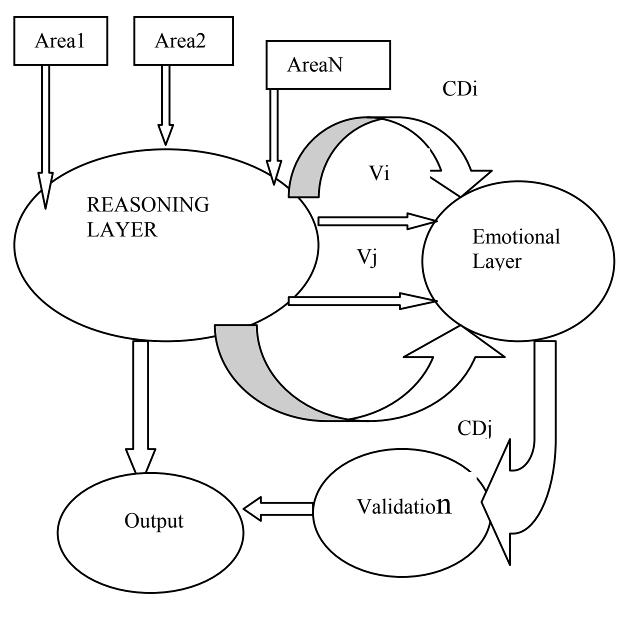

Fig. 4 Global Network

and the consequent activations are quite different. Notice that the output of this layer is to be interpreted again as the activation ratio of its neuron outputs. Then, the result of the reasoning layer is validated by the activation ratio of the emotional layer. On fig. 4 the overall topology of the network is depicted.

IV CONCLUSIONS

The capability of fuzzy neural networks to perform relational operations, in addition to the qualitative reasoning, is here used to implement an emotional layer, which function is to validate the output of the reasoning layer. This means that no consequence is taken from the antecedent inputs if not enabled by the emotional layer. Another possibility is to impose a consequent obtained from some rules, discarding (on the implicit defuzzification process) the results obtained from other rules. In this paper the first option has been taken.

Two different problems (although linked between themselves) may arise, in order to make necessary the extra layer: there is not enough learning about some rules, or, a real experiment may simultaneously activate (with relatively high intensity) different rules with contradictory consequent results – which may originate an undesirable and dangerous “averaging” effect.

Since each reasoning layer neuron gets its own information about each individual rule, specifically: one bit stating if there has been or not teaching of that rule and the other bit indicating the taught value. Using each and every neuron addressed rule information (one and only one rule per neuron is addressed) it has been developed an emotional layer that uses that information as inputs and gives an output activation ratio that inhibits or not the consequent result of the reasoning layer.

REFERENCES

[1] Damásio, A. (1995). Descartes’Error: Emotion, Reason and the Human Brain. Avon Edition.

[2] Gupta, M. and QI "On fuzzy neurone models" Proc. Int. Joint Conf. Neural Networks, volII, 431-436, Seattle. (1991).

[3] Horikawa, S., T. Furuhashi and Y. Uchikawa "On fuzzy modelling using fuzzy neural networks with the back propagation algorithm." IEEE Transactions Neural Networks 3(5): 801-806. (1992).

[4] Keller,J.M. and D.J.Hunt"Incorporating fuzzy membership functions into the perceptron algorithm." IEEE Transactions Pattern Anal.Mach.Intell.. PAMI-7(6):693-699. (1985).

[5] Lin, Chin-Ten and Lee,C.S. A Neuro-Fuzzy Synergism to Intelligent Systems.New Jersey : Prentice Hall. (1996).

[6] Parzen, E.(1962). "On Estimation of a probability density function and mode" *Ann. Math. Stat.*, 33, 1065-1072.

[7] Pedrycz,W. "Fuzzy neural networks with reference neurones as pattern classifiers" IEEE Trans.Neural Networks 3(5):770-775. (1992).

[8] Tomé,J.A.."Neural Activation ratio based Fuzzy Reasoning."Proc. IEEE World Congress on Computational Inteligence, Anchorage,May 1988,pp 1217-1222.

[9] Tomé,J.A."Counting Boolean Networks are Universal Approximators" Proc. of 1998 Conference of NAFIPS, Florida, August 1988, pp 212-216.

[10] Tomé,J.A. and Carvalho, J.P. "Rule Capacity in Fuzzy Boolean Networks”. "Proc. of NAFIPS-FLINT 2002 International Conference, New Orleans, IEEE2002, pp 124-128.

[11] Tomé,J.A. "Fuzzy Evaluation of Relational Operators in Boolean Nets."Proc. of EUROFUSE-SIC99, Budapest, May, 1999, pp. 218-222. (1994).

[12] Tomé, J.A. “Learning in Fuzzy Boolean Nets-Rule Distinguishing Power”. Proc. of the SBIA 2002, Porto Galinhas, Recife