Letters

Comments on "Functional Equivalence Between Radial Basis Function Networks and Fuzzy Inference Systems"

H. C. Andersen, A. Lotfi, and L. C. Westphal

Abstract—**The above paper claims that under a set of minor restrictions radial basis function networks and fuzzy inference systems are functionally equivalent. The purpose of this letter is to show that this set of restrictions is incomplete and that, when it is completed, the said functional equivalence applies only to a small range of fuzzy inference systems. In addition, a modified set of restrictions is proposed which is applicable for a much wider range of fuzzy inference systems.**

Index Terms—**Artificial neural networks, functional equivalence, fuzzy inference systems, radial basis function networks.**

I. Introduction

In the above paper[1] Jang and Sun stated that radial basis function networks (RBFN's) and fuzzy inference systems (FIS's) are functionally equivalent under some apparently minor restrictions. They stated this equivalence as follows:

"... the functional equivalence between an RBFN and a fuzzy inference system can be established if

1) The number of receptive field units is equal to the number of fuzzy if–then rules.
2) The output of each fuzzy if–then rule is composed of a constant.
3) The membership functions within each rule are chosen as Gaussian functions with the same variance.
4) The T-norm operator used to compute each rule's firing strength is multiplication.
5) Both the RBFN and the fuzzy inference system under consideration use the same method (i.e., either weighted average or weighted sum) to derive their overall outputs."

It will be shown that an additional necessary but unmentioned restriction on the stated functional equivalence is that the number of dependencies of rules upon membership functions is limited to one; this restriction will be called the *unary dependency restriction.* This is demonstrated by use of an example in which the above five restrictions are all satisfied but where the equivalence breaks down. It is then argued that with application of the unary dependency restriction the functional equivalence becomes so restrictive as to be unusable for almost all FIS's. Finally, a new set of conditions required for functional equivalence is suggested which does not include the unary dependency restriction and which is consequently more widely applicable.

Manuscript received August 5, 1997.

H. C. Andersen and L. C. Westphal are with the Department of Electrical and Computer Engineering, University of Queensland, Brisbane, Queensland, 4072, Australia.

A. Lofti is with the Department of Manufacturing Engineering, The Nottingham Trent University, Nottingham, U.K.

Publisher Item Identifier S 1045-9227(98)08571-3.

[1] J.-S. R. Jang and C.-T. Sun, *IEEE Trans. Neural Networks*, vol. 4, pp. 156–159, Jan. 1993.

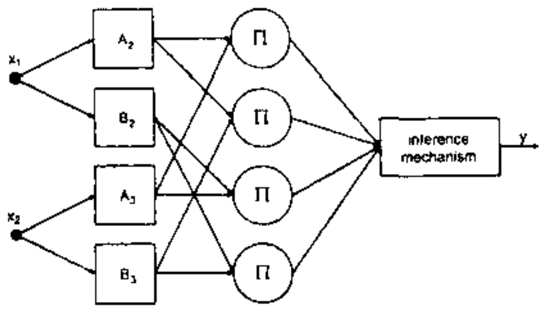

Fig. 1. FIS with two inputs, four membership functions, and four rules. The membership functions are all utilized by more than one rule. This FIS is of a verycommon type but it does not satisfy the unary dependency restriction.

II. Analysis of Functional Equivalence

The following example will illustrate a case where first a simple FIS is designed; this FIS is converted to its functionally equivalent RBFN (satisfying the five restrictions quoted above); a learning algorithm adapts the RBFN; and then an attempt is made to convert the RBFN back to its original FIS structure.

Example Consider a FIS (see Fig. 1) with two inputs, x_1 and x_2, and one output, y. For each input two fuzzy sets, A_1, A_2, and B_1, B_2, are defined. The membership functions for the fuzzy sets are

$$\mu_{A_1}(x_1) = \exp\left(-\frac{(x_1 - c_{A_1})^2}{\sigma^2}\right)$$
$$\mu_{A_2}(x_1) = \exp\left(-\frac{(x_1 - c_{A_2})^2}{\sigma^2}\right)$$
$$\mu_{B_1}(x_2) = \exp\left(-\frac{(x_2 - c_{B_1})^2}{\sigma^2}\right)$$
$$\mu_{B_2}(x_2) = \exp\left(-\frac{(x_2 - c_{B_2})^2}{\sigma^2}\right).$$

There are a maximum of four rules which can be considered for this FIS; these are

rule 1: if x_1 is A_1 and x_2 is B_1, then y is f_1
rule 2: if x_1 is A_1 and x_2 is B_2, then y is f_2
rule 3: if x_1 is A_2 and x_2 is B_1, then y is f_3
rule 4: if x_1 is A_2 and x_2 is B_2, then y is f_4.

The firing strength, $w_i : i = 1 \cdots 4$, for each rule is calculated as follows:

$$w_1 = \mu_{A_1}(x_1)\mu_{B_1}(x_2)$$
$$w_2 = \mu_{A_1}(x_1)\mu_{B_2}(x_2)$$
$$w_3 = \mu_{A_2}(x_1)\mu_{B_1}(x_2)$$
$$w_4 = \mu_{A_2}(x_1)\mu_{B_2}(x_2).$$

The inference mechanism used to compute the crisp output, y, is of the weighted average form, that is

$$y = \frac{\sum_{i=1}^4 f_i w_i}{\sum_{i=1}^4 w_i}.$$

It is noted that, as a consequence of restriction 3) and the use of each membership function by more than one rule, all of the membership functions must have the same variance. It is argued that this is an unusual condition because FIS's usually permit each membership functions to be set independently. Nevertheless, this FIS satisfies all of the restrictions quoted from [1] above and we can proceed to convert it to its equivalent RBFN.

To obtain a RBFN which should, according to the quoted restrictions, be equivalent to the above FIS, we required four receptive field units and an output which is calculated by the weighted average method. This can be described as follows:

$$w_1 = \exp\left(-\frac{(x_1-\tilde{c}_{11})^2+(x_2-\tilde{c}_{12})^2}{\tilde{\sigma}_1^2}\right)$$
$$w_2 = \exp\left(-\frac{(x_1-\tilde{c}_{21})^2+(x_2-\tilde{c}_{22})^2}{\tilde{\sigma}_2^2}\right)$$
$$w_3 = \exp\left(-\frac{(x_1-\tilde{c}_{31})^2+(x_2-\tilde{c}_{32})^2}{\tilde{\sigma}_3^2}\right)$$
$$w_4 = \exp\left(-\frac{(x_1-\tilde{c}_{41})^2+(x_2-\tilde{c}_{42})^2}{\tilde{\sigma}_4^2}\right)$$
$$y = \frac{\sum_{i=1}^4 f_i w_i}{\sum_{i=1}^4 w_i}$$

where the parameters, $\tilde{c}_{ij}$ and $\tilde{\sigma}_i$ for $i = 1, 2, 3, 4$ and $j = 1, 2$, are given as follows:

$$\tilde{c}_{11} = c_{A_1}, \quad \tilde{c}_{21} = c_{A_1}, \quad \tilde{c}_{31} = c_{A_2}, \quad \tilde{c}_{41} = c_{A_2},$$
$$\tilde{c}_{12} = c_{B_1}, \quad \tilde{c}_{22} = c_{B_2}, \quad \tilde{c}_{32} = c_{B_1}, \quad \tilde{c}_{42} = c_{B_2}.$$

$$\tilde{\sigma}_1 = \tilde{\sigma}_2 = \tilde{\sigma}_3 = \tilde{\sigma}_4 = \sigma$$

Notice that many of the parameters of the RBFN are set to the same value, e.g. $\tilde{c}_{11} = \tilde{c}_{21} = c_{A_1}$. Fig. 2 demonstrates how this forces the centres of the radial basis functions (RBF's) to be alligned in a *grid formation*. This arrangement forces the RBFs' centres to be aligned in the input dimensions in which they share membership functions.

Now a learning algorithm is applied to this RBFN (as is suggested in [1]) which results in the parameters' being changed as follows:

$$\tilde{c}_{ij} \rightarrow \tilde{c}'_{ij} \quad \text{for } i = 1, 2, 3, 4 \quad \text{and} \quad j = 1, 2$$
$$\tilde{\sigma}_i \rightarrow \tilde{\sigma}'_i \quad \text{for } i = 1, 2, 3, 4.$$

When an attempt is made to convert the RBFN back to a FIS a problem arises. The parameters of the RBFN have all changed and since there is not a one-to-one correspondence between the parameters of the RBFN and the parameters of the FIS, a functionally equivalent conversion back to the original FIS cannot be made. The simplest FIS which would be equivalent to the adapted RBFN would have eight membership functions (see Fig. 3), not four like the original FIS. Indeed if, for instance, the original membership function A_1 had defined a fuzzy set *hot*, then the FIS equivalent of the adapted RBFN would have two definitions for *hot* on the same crisp input; this would generally be considered confusing.

Comments: It is clear from the above example that the functional equivalence as described by Jang and Sun [1] is only valid when each rule has a separate set of membership functions or, in other words, each membership function is used by **at most one rule**. This implies that an additional necessary restriction—the unary dependency restriction—on the functional equivalence would be as follows.

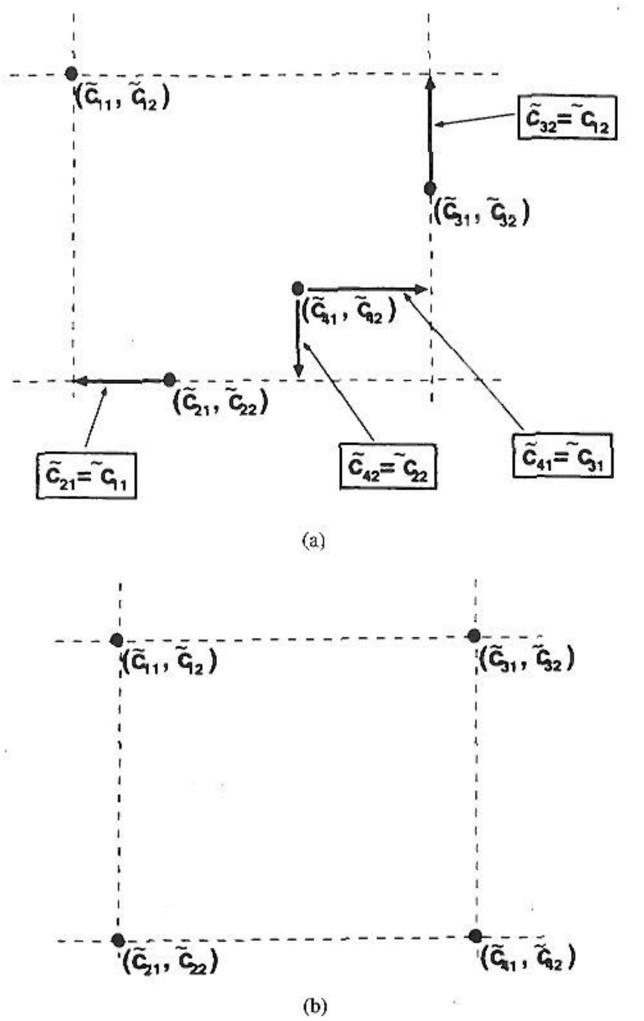

Fig. 2. The centers of four RBF's plotted in two dimensions. (a) Shows four centers which do not satisfy the constraints imposed by the sharing of membership functions by rules, but have arrows pointing to the positions which the constraints (marked in boxes) force them toward. (b) Shows the positions of the centers once the constrains have been satisfied.

Each membership function can be used by only one rule.

This makes it clear that the range of applicability of the original functional equivalence definition is much narrower than was implied by the original restrictions since most FIS's utilize each membership function more than once. For example, it is common to have combinations of rules like "*if temperature is hot and humidity is dry then...*" and "*if temperature is hot and humidity is wet then...*" in which the fuzzy set, *hot*, is utilized by more than one rule.

The problem described above is a quite significant disadvantage, however it is possible to restate the functional equivalence definition slightly and thereby make it applicable to a much wider range of FIS's.

III. Alternative Statement of Functional Equivalence

Firstly, as stated above, rules which share a membership function must be aligned in the input-dimension on which that function is defined. In the general case this results in rules being placed on intersections of a multidimensional grid. The RBF's in a RBFN must maintain this relationship. This may be done by constraining the positions of the RBF's to a multidimensional grid formation and constraining the variances of all RBF's on common grid-lines to have

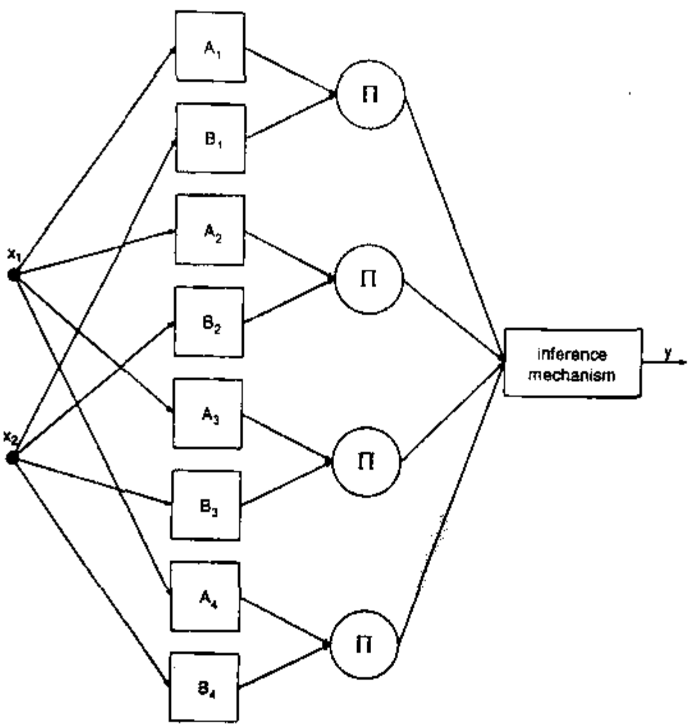

Fig. 3. FIS with two inputs, eight membership functions, and four rules. Each of the membership functions is utilized by only one rule. This FIS satisfies the unary dependency restriction, but is of an unusual type.

the same input variances; only one membership function needs to be defined for each grid-line.

Secondly, by using a RBF which has a separate variance for each input the interdependency of the variances is removed, i.e. each membership function in the FIS can have a different Gaussian membership function; this type of RBF can be written as

$$w_i = \exp\left(-\sum_{j=1}^{N} \frac{(x_j - c_{ij})^2}{\sigma_{ij}^2}\right)$$

where N is the number of inputs.

Combining these two considerations with the restrictions as originally stated we obtain the following more widely applicable definition for the functional equivalence between FIS's and RBFN's.

The functional equivalence between an RBFN and a fuzzy inference system can be established if the following restrictions hold.

1) The number of receptive field units in the RBFN is equal to the number of fuzzy if–then rules.
2) The output of each fuzzy if–then rule is composed of a constant.
3) The membership functions within each rule are chosen as Gaussian functions.
4) The T-norm operator used to compute each rule's firing strength is multiplication.
5) Both the RBFN and the FIS under consideration use the same method (i.e., either weighted average or weighted sum) to derive their overall outputs.
6) The positions of the RBF's are on a multidimensional grid and RBF's on common grid-lines have the same input-variances.
7) Each RBF has a separate variance for each input.

Restriction 6) implies that RBF's initially located on particular grid-lines (as defined by membership functions) must remain on those grid-lines; these grid-lines may however be moved which is equivalent to changing the parameters of the corresponding membership functions.

IV. Conclusion

In this letter we have described how the functional equivalence between RBFN's and FIS's, as stated in [1], effectively applies only to FIS's which have a separate set of fuzzy sets for each rule. It is argued that this is very restrictive and, in order to make the result more generally applicable, a new definition for the functional equivalence between RBFN's and FIS's is proposed.

Author's Reply

Jyh-Shing Roger Jang

In our original paper mentioned above[1] on functional equivalence between radial basis function networks and fuzzy inference systems, we were emphasizing the equivalence in terms of "forward calculation," or the equivalence in input–output function. We did not address the functional equivalence throughout a training process due to the following reason.

Training/learning is a complicated process which may employ various stochastic/deterministic optimization schemes, such as steepest descent (or backpropagation), Gauss–Newton method, Levenberg–Marquardt method, genetic algorithms, simulated annealing, Tabu search, downhill simplex method, etc. Each training method itself also involves the tuning/updating of many control parameters. Therefore instead of talking about functional equivalence throughout a possibly complicated/diversified training method, we decided to make it simple and confined our discussion to the equivalence in model forward calculation. Once the functional equivalence is clearly established, its counterpart throughout a training process can be inferred directly based on the specific training method used.

Admittedly, the original publication uses a grid partition for a fuzzy inference system, which might break the functional equivalence with a radial basis function network during training. However, it should be straightforward to use a scatter partition instead (in which each membership function is only used in a rule), as suggested by Jang's other publication [2]. The comments by Anderson *et al.* provide a detailed coverage of this issue. Moreover, an in-depth discussion of the functional equivalence issue can be found in [1].

In fact, at the writing of the original paper on functional equivalence, the research on neuro-fuzzy literature was still young and the purpose of that paper was to remind people that there is a strong tie between two major paradigms in the realms of artificial neural networks and fuzzy systems. Today we have a better understanding of those "local models," which include:

- radial basis function networks;

Manuscript received May 14, 1998.
The author is with the Computer Science Department, National Tsing Hua University, Hsinchu, Taiwan.
Publisher Item Identifier S 1045-9227(98)06179-7.

[1] J.-S. R. Jang and C.-T. Sun, *IEEE Trans. Neural Networks,* vol. 4, pp. 156–159, Jan. 1993.

- fuzzy inference systems;
- modular networks (mixture of experts);
- generalized regression neural networks [3];
- Nadaraya–Watson kernel regression.

These models might come from different disciplines, but they can all be made functionally equivalent under certain conditions.

References

[1] K. J. Hunt, R. Hass, and R. Murray-Smith, "Extending the functional equivalence of radial basis function networks and fuzzy inference systems," *IEEE Trans. Neural Networks*, vol. 7, pp. 776–781, May 1996.

[2] J.-S. R. Jang, C.-T. Sun, and E. Mizutani, *Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence*, MATLAB Curriculum Series. Upper Saddle River, NJ: Prentice-Hall, 1997.

[3] D. F. Specht, "A generalized regression neural network," *IEEE Trans. Neural Networks*, vol. 2, pp. 568–576, Nov. 1991.

Errata to "Model Transitions in Descending FLVQ"

Andrea Baraldi, Palma Blonda, Flavio Parmiggiani, Guido Pasquariello, and Giuseppe Satalino

In the above paper,[1] the names of two authors were exchanged in the byline and biographies. The correct names are Guiseppe Satalino and Guido Pasquariello.

Manuscript received September 30, 1998. A. Baraldi and F. Parmiggiani are with IMGA_CNR, Bologna 40129, Italy.

P. Blonda, G. Pasquariello, and G. Satalino are with IESI-CNR, 70126 Bari, Italy.

Publisher Item Identifier S 1045-9227(98).

[1] A. Baraldi, P. Blonda, F. Parmiggiani, G. Pasquariello, and G. Satalino, *IEEE Trans. Neural Networks*, vol. 9, pp. 724–738, Sept. 1998.