1

Functional Equivalence between Radial Basis Function Networks and Fuzzy Inference Systems

J.-S. Roger Jang and C.-T Sun
Department of Electrical Engineering and Computer Science
University of California, Berkeley, CA 94720
jang@diva.berkeley.edu

Abstract

This short article shows that under some minor restrictions, the functional behavior of radial basis function networks and fuzzy inference systems are actually equivalent. This functional equivalence implies that advances in each literature, such as new learning rules or analysis on representational power, etc., can be applied to both models directly. It is of interest to observe that two models stemming from different origins turn out to be functional equivalent.

I. Introduction

This paper demonstrates the functional equivalence between radial basis function networks (RBFN's) and a simplied class of fuzzy inference systems. Though these two models are motivated from different origins (RBFN's from physiology and fuzzy inference systems from cognitive science), they share common characteristics not only in their operations on data, but also in their learning process to achieve desired mappings. We show that under some minor restrictions, they are functionally equivalent; the learning algorithms and the analysis on representational power for one model can be applied to the other, and vice versa.

II. Radial Basis Function Networks

The locally-tuned and overlapping receptive field is a well-known structure that has been studied in regions of cerebral cortex, the visual cortex, etc. Based on the biological receptive

*Research supported in part by NASA Grant NCC 2-275, MICRO Grant 92-180 and EPRI Agreement RP 8010-34.

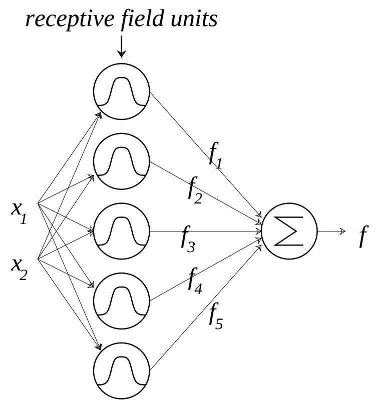

Figure 1: *A radial basis function network (RBFN)* .

fields, Moody and Darken [7, 8] proposed a network structure called *radial basis function network* (RBFN) which employs local receptive fields to perform function mappings. Figure 1 shows the schematic diagram of an RBFN with five receptive field units; the output of i-th receptive field unit (or hidden unit) is

$$w_i = R_i(\vec{x}) = R_i(\|\vec{x} - \vec{c_i}\|/\sigma_i), \; i = 1, 2, ..., H \tag{1}$$

where $\vec{x}$ is an N-dimensional input vector, $\vec{c_i}$ is a vector with the same dimension as $\vec{x}$, H is the number of receptive field units, and $R_i(\cdot)$ is the i-th receptive field response with a single maximum at the origin. Typically, $R_i(\cdot)$ is chosen as a Gaussian function

$$R_i(\vec{x}) = exp[-\frac{\|\vec{x} - \vec{c_i}\|^2}{\sigma_i^2}]. \tag{2}$$

Thus the radial basis function w_i computed by the i-th hidden units is maximum when the input vector $\vec{x}$ is near the center $\vec{c_i}$ of that unit.

The output of a radial basis function networks can be computed in two ways. For the simpler one, as shown in Figure 1, the output is the weighted sum of the function value associated with each receptive field:

$$f(\vec{x}) = \sum_{i=1}^{H} f_i w_i = \sum_{i=1}^{H} f_i R_i(\vec{x}), \tag{3}$$

where f_i is the function value, or strength, of i-th receptive field. With the addition of lateral connections (not shown in Figure 1) between the receptive field units, the network can produce the normalized response function as the weighted average of the strengths [7]:

$$f(\vec{x}) = \frac{\sum_{i=1}^{H} f_i w_i}{\sum_{i=1}^{H} w_i} = \frac{\sum_{i=1}^{H} f_i R_i(\vec{x})}{\sum_{i=1}^{H} R_i(\vec{x})}. \tag{4}$$

To minimize the square errors between desired output and model output, several learning algorithms have been proposed to identify the parameters ($\vec{c_i}$, σ_i and f_i) of an RBFN. Moody et al. [7] use a self-organizing techniques to find the centers ($\vec{c_i}$) and widths (σ_i) of the receptive fields, and then employ the supervised Adaline or LMS learning rule to identify f_i. On the other hand, Chen et al. [1] apply orthogonal least squares learning algorithm to determine those parameters. Other types of learning algorithms and variant structures can be found in [6, 9].

III. Fuzzy If-Then Rules and Fuzzy Inference Systems

An example of *Fuzzy if-then rules* (or *fuzzy conditional statement*) is

If pressure is high, then volume is small.

where *pressure* and *volume* are *linguistic variables* [16], *high* and *small* are *linguistic values* (or *linguistic labels*) characterized by appropriate membership functions. Another type of fuzzy if-then rule, proposed by Takagi and Sugeno [12], has fuzzy sets involved only in the premise part. For instance, the dependency of the air resistance (force) on the speed of a moving object can be described as

$$\textit{If velocity is high, then } force \;=\; k * (velocity)^2.$$

where *high* is the only linguistic label here, and the consequent part is described by a nonfuzzy equations of the input variable, velocity.

Fuzzy inference systems are also known as *fuzzy rule based systems*, *fuzzy models*, *fuzzy associative memories*, or *fuzzy controllers* when used as controllers. A fuzzy inference system is composed of a set of fuzzy if-then rules, a database containing membership functions of linguistic labels, and and an inference mechanism called *fuzzy reasoning*. Suppose we have a rule base consisting of two fuzzy if-then rules of Takagi and Sugeno's type:

$$\text{Rule 1: If } x_1 \text{ is } A_1 \text{ and } x_2 \text{ is } B_1, \text{ then } f_1 = a_1x_1 + b_1x_2 + c_1,$$
$$\text{Rule 2: If } x_1 \text{ is } A_2 \text{ and } x_2 \text{ is } B_2, \text{ then } f_2 = a_2x_1 + b_2x_2 + c_2,$$

then the fuzzy reasoning mechanism can be illustrated in Figure 2 (a) where the firing strength (or weight) of i-th rule is obtained as the T-norm (usually multiplication or min. operator) of the membership values on the premise part

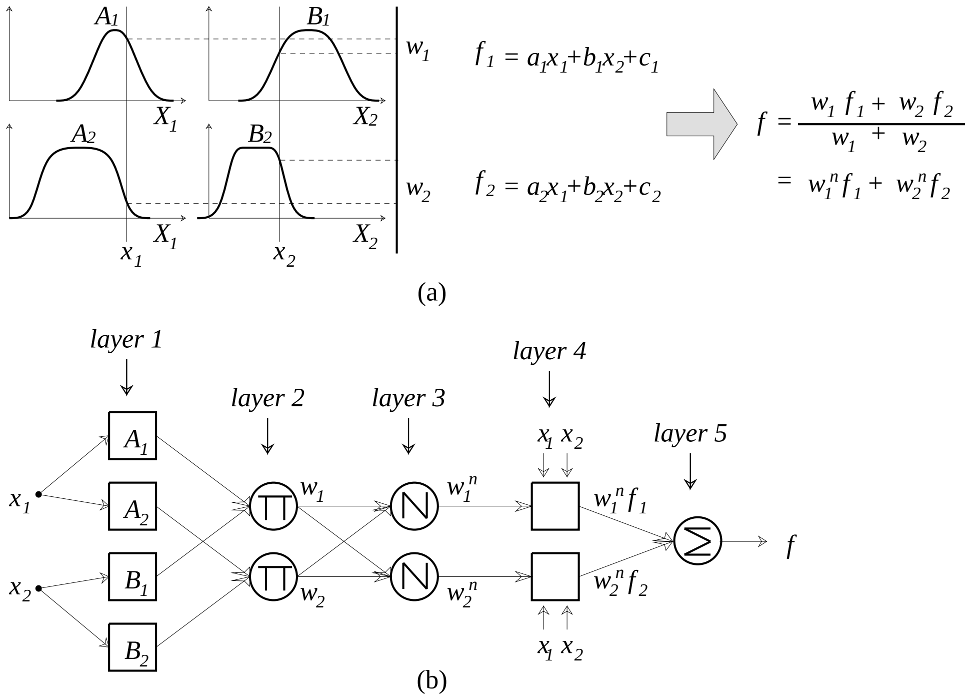

Figure 2: *(a) fuzzy reasoning; (b) adaptive network representation*

$$
\begin{aligned}
w_i &= \mu_{A_i}(x_1)\mu_{B_i}(x_2),\ or \\
&= min\{\mu_{A_i}(x_1), \mu_{B_i}(x_2)\}.
\end{aligned} \tag{5}
$$

Note that the overall output can be chosen either as the weighted sum of each rule's output [11, 2]

$$f(\vec{x}) = \sum_{i=1}^{R} w_i f_i, \tag{6}$$

or more conventionally, as the weighted average [12] (as shown in Figure 2 (a))

$$f(\vec{x}) = \frac{\sum_{i=1}^{R} w_i f_i}{\sum_{i=1}^{R} w_i}, \tag{7}$$

where R is the number of fuzzy if-then rules.

Fuzzy modeling concerns the identification of the structure (number of rules, partition pattern, etc.) and parameters of fuzzy inference systems. Various methodologies of fuzzy modeling have been proposed in the past years. Takagi et al. [13] and Sugeno et al. [10] employ nonlinear programming and heuristic search to identify both the structure and parameters. Hariawa et al. [2] and Takagi et al. [11] introduce feedforward neural networks into fuzzy inference systems and solve the parameter identification problems through neural network's learning algorithm. We [4, 3, 5] propose a more direct method which transforms the fuzzy inference system into an adaptive network (Figure 2 (b)) and then employ both the back-propagation-type gradient descent to update premise parameters (which determine the shapes and positions of membership functions) and the least square method to identify consequent parameters (which specify the output of each rule). In the proposed adaptive network shown in Figure 2 (b), there is no weight associated with each link and nodes in different layer can have different functions corresponding to each steps in the fuzzy reasoning mechanism. More specifically, layer 1 calculates membership values, layer 2 perform T-norm operator (multiplication in this case), layer 3 computes normalized weights, layer 4 derives the product of each rule's output and corresponding normalized weight, layer 5 sums its inputs as the overall output. Note that layer 1 contains premise parameters and layer 4 contains consequent parameters; all the other layers perform fixed functions without modifiable parameters. For an in-depth treatment, see [5].

IV. Functional Equivalence and Its Implication

From equation (3), (4) and equation (6), (7), it is obvious that 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. (Namely, a_1, b_1, a_2 and b_2 are zeros in Figure 2 (a).)
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.

Under these conditions, the membership functions of linguistic labels A_1 and B_1 in Figure 2 (a) can be expressed as

$$\mu_{A_1}(x_1) = exp[-\frac{(x_1 - c_{A_1})^2}{\sigma_1^2}], \;\; \mu_{B_1}(x_2) = exp[-\frac{(x_2 - c_{B_1})^2}{\sigma_1^2}]. \tag{8}$$

Hence the firing strength (or weight) of rule 1 (the output of the first node in layer 2) is

$$w_1(x_1, x_2) = \mu_{A_1}(x_1)\mu_{B_1}(x_2) = exp[-\frac{\|\vec{x} - \vec{c_1}\|^2}{\sigma_1^2}] = R_i(\vec{x}), \tag{9}$$

where $\vec{c_1} = (c_{A_1}, c_{B_1})$, the center of the corresponding receptive field. The same argument applies to w_2. Therefore under the above constraints, the output of Figure 2(a) or (b) is exactly the same as an RBFN (with two receptive field units) where the receptive field units and output units are functionally equivalent to the cascades of layer 1, 2 and layer 3, 4, 5, respectively, in Figure 2. Without the above constraints, RBFN's are only a special case of fuzzy inference systems.

Because of this functional equivalence, we can apply advances and new developments of one model to the other, and vice versa. In other words, we can apply the learning rules of RBFN's mentioned in section II to fuzzy inference systems, and the learning rules of fuzzy inference systems in section III can also be utilized to find the structure (i.e., number of receptive field units) and parameters of RBFN's. Moreover, recently Wang [14, 15] proved that a fuzzy inference system with membership functions of scaled Gaussian functions

$$\mu_A(x) = k * exp[-\frac{(x-c)^2}{\sigma^2}] \tag{10}$$

is actually a universal approximator that can approximate any input-output data arbitrarily well on a compact set. This argument can be readily applied to RBFN's if the receptive field response in equation (2) is also scaled by a constant.

V. Concluding Remarks

In this letter, we briefly introduce the structure and learning rules of radial basis function networks (RBFN's) and fuzzy inference systems. Some minor restrictions that renders the functional equivalence of these two models are also discussed. Due to the equivalence of these models, it becomes straightforward to apply one model's learning rules to the other, and vice versa. Furthermore, we can claim both models are universal approximators if the receptive field responses and membership functions are chosen as a scaled version of Gaussian functions. It is of interest to observe that these two models, though derived from different origins and each with different interpretations on their process of data, turn out to be functionally equivalent.

Acknowledgement

The guidance and help of Professor Lotfi A. Zadeh and other members of the "fuzzy group" at UC Berkeley is gratefully acknowledged.

References

[1] S. Chen, C. F. N. Cowan, and P. M. Grant. Orthogonal least squares learning algorithm for radial basis function networks. *IEEE Trans. on Neural Networks*, 2(2):302–309, March 1991.

[2] S. Horiawa, T. Furuhashi, S. Ouma, and Y. Uchikawa. A fuzzy controller using a neural network and its capability to learn expert's control rules. In *Proc. of the International Conference on Fuzzy Logic and Neural Networks*, pages 103–106, Japan, 1990.

[3] J.-S. Roger Jang. Fuzzy modeling using generalized neural networks and Kalman filter algorithm. In *Proc. of the Ninth National Conference on Artificial Intelligence (AAAI-91)*, pages 762–767, July 1991.

[4] J.-S. Roger Jang. Rule extraction using generalized neural networks. In *Proc. of the 4th IFSA World Congress*, July 1991.

[5] J.-S. Roger Jang. ANFIS: Adaptive-network-based fuzzy inference systems. *IEEE Trans. on Systems, Man, and Cybernetics*, 1992. (Forthcoming).

[6] R. D. Jones, Y. C. Lee, C. W. Barnes, G. W. Flake, K. Lee, and P. S. Lewis. Function approximation and time series prediction with neural networks. In *Proc. of IEEE International Joint Conference on Neural Networks*, pages I–649–665, 1990.

[7] J. Moody and C. Darken. Learning with localized receptive fields. In D. Touretzky, G. Hinton, and T. Sejnowski, editors, *Proc. of the 1988 Connectionist Models Summer School*. Carnegie Mellon University, Morgan Kaufmann Publishers, 1988.

[8] J. Moody and C. Darken. Fast learning in networks of locally-tuned processing units. *Neural Computation*, 1:281–294, 1989.

[9] M. T. Musavi, W. Ahmed, K. H. Chan, K. B. Faris, and D. M. Hummels. On the training of radial basis function classifiers. *Neural Networks*, 5(4):595–603, 1992.

[10] M. Sugeno and G. T. Kang. Structure identification of fuzzy model. *Fuzzy Sets and Systems*, 28:15–33, 1988.

[11] H. Takagi and I. Hayashi. NN-driven fuzzy reasoning. *International Journal of Approximate Reasoning*, 5(3):191–212, 1991.

[12] T. Takagi and M. Sugeno. Derivation of fuzzy control rules from human operator's control actions. *Proc. of the IFAC Symp. on Fuzzy Information, Knowledge Representation and Decision Analysis*, pages 55–60, July 1983.

[13] T. Takagi and M. Sugeno. Fuzzy identification of systems and its applications to modeling and control. *IEEE Trans. on Systems, Man, and Cybernetics*, 15:116–132, 1985.

[14] L.-X. Wang. Fuzzy systems are universal approximators. In *Proc. of the IEEE International Conference on Fuzzy Systems*, San Diego, March 1992.

[15] L.-X. Wang and J. M. Mendel. Fuzzy basis function, universal approximation, and orthogonal least squares learning. *IEEE Trans. on Neural Networks*, September 1992. (Forthcoming).

[16] L. A. Zadeh. Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Trans. on Systems, Man, and Cybernetics*, 3(1):28–44, January 1973.