HIERARCHICAL FUZZY SYSTEMS

Radek Šindelář *,1

* *Center for Applied Cybernetics*
Czech Technical University
Karlovo náměstí13/E, 131 25 Prague, Czech Republic

Abstract: A method for a fuzzy hierarchical structure design is presented. The proposed method uses data to design a structure of the fuzzy subsystems. The fuzzy structure is designed level by level from data thus developing an initial fuzzy model is avoided. The methods is tested on one real-world application - the daily gas consumption prediction.

Keywords: fuzzy modelling, hierarchical fuzzy system, rule base

1. INTRODUCTION

A rule explosion is a fundamental limitation of fuzzy systems because the number of rules increases exponentially as the number of input variables increases. Suppose that we have n inputs and m fuzzy sets defined for each of them then the number of the rules of the standard fuzzy system is m^n. A rule base with many input variables and the huge number of rules tends to lose all good features - transparency, ability to generalize, accuracy etc. Hierarchical organization of fuzzy rule bases is the way how to reduce the complexity of the fuzzy system and improve the insight into the system behaviour. Also design, transparency, tuning etc. become easier for the system consisting of smaller fuzzy systems.

The idea of using hierarchical fuzzy systems has been reported by many authors (Wang, 1998; Gegov and Frank, 1995; Joo, 2003; Wang, 1999). In (Wang, 1998) it is proved that the hierarchical fuzzy system is an universal approximator and the number of rules in a hierarchical system is linearly proportional to the number of input variables. For the structure depicted in Fig. 7 it can be shown that the number of rules is minimal.

[1] This work was in part supported by the Ministry of Education of the Czech Republic under Project 1M6840770004.

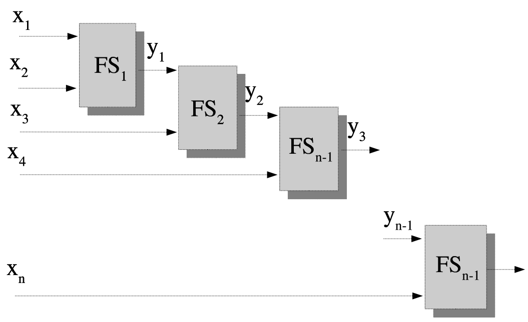

Fig. 1. An example of a hierarchical system with n-inputs which consists of two-input fuzzy systems

Several approaches lead to the design of the hierarchical fuzzy system. Usually the large fuzzy system is decomposed into several fuzzy systems whose structure is simple see (Joo, 2003; Gegov and Frank, 1995). The small fuzzy systems are interconnected according to the given topology. This approach still need to develop a large complex fuzzy system. The performance and generalization of this model can be poor thus the splitting process should fail. Usually some kind of the rule base analysis is used to split the large rule base. In (Wang, 1999) Wang proposes the learning gradient descent algorithm for designing the hierarchical fuzzy system from the input-output data. But some assumptions on the inner structure of the subsystems (number and positions of fuzzy membership functions) are necessary concerning also inner variables. Our goal is to create a hierarchical fuzzy system layer by layer from the data in order to avoid developing the large fuzzy system without using prior information about internal structures of fuzzy subsystems.

The paper is organized as follows: the structure of fuzzy hierarchical systems is discussed in the next section. The algorithm of the fuzzy hierarchical system design is presented in following sections. The section 3 presents the example and the basic analysis. The problem of inner fuzzy connections is discussed in the section 4. After creating the hierarchical structure, the rule bases can be simplified. This step is described in the section 5. The real-world example is shown in the section 6.

2. STRUCTURE OF HIERARCHICAL FUZZY SYSTEMS

In general the hierarchical fuzzy system consists of more then one rule base where an output of one rule base serves as an input to another rule base. Let denote inputs of subsystems U, U consists of real inputs x_i or inner connections z_i.

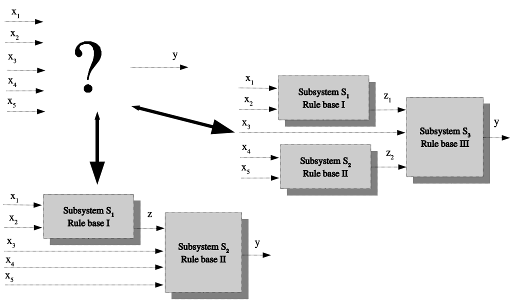

Fig. 2. The structure design

Choosing internal structure of fuzzy systems is crucial. Usually there is a set of available inputs and the aim is to propose a suitable structure (Fig. 2). Despite the fact that the structure in Fig. 7 has the minimal number of rules it is not suitable for modelling real systems. Such a structure is not well-arranged with respect to the physical meaning of simple fuzzy subsystems. Thus the better way is to define the structure of the fuzzy system based on knowledge of the modelled system. Inputs with the same background or related to the same output should be connected to the same fuzzy subsystem. It is supposed in this work that the internal structure is given at the beginning.

Passing information from the rule base I to the rule base IV in Fig. 3 can be realized either using defuzzification of the output of the rule base I and subsequent fuzzification at the input of the rule base III or passing fuzzy sets from the output of the rule base I at the input of the rule base IV directly. Defuzzification and fuzzification need to define fuzzy membership functions for variable z_1 and it is more time consuming than passing fuzzy sets. If there are no data set for z_1 then the properties of the variable z_1 can not be verified. This method should be used if it is necessary to get crisp values at the output of the rule base I, e.g. z_1 is a state of the dynamic system or it is one of the crisp outputs of the hierarchical fuzzy system. Passing fuzzy sets has some advantages comparing to the previous approach. No special knowledge about inner variables are required.

3. DESIGN OF HIERARCHICAL FUZZY SYSTEMS FROM DATA

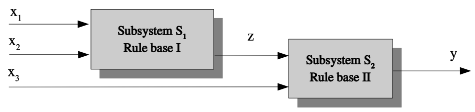

Fig. 3. The simple structure of the hierarchical fuzzy system

Suppose the simple structure of the hierarchical fuzzy system according to the Fig. 3. The fuzzy system has three input variables x_1, x_2, x_3 and one output y. It consists of two subsystems S_1, S_2 having the rule base I and the rule base II . The output z of the subsystem S_1 is used as an input to the subsystem S_2. We would like to create the hierarchical fuzzy system using fuzzy clustering algorithms from the measured data set. The algorithm starts to create the subsystems at the lower level, it means the subsystem which inputs are the real (measured) inputs. Then the results from the first level are passed as the inputs to the next one. In conventional fuzzy hierarchical systems z is a fuzzy variable described by a fuzzy set. The rule base of the fuzzy system S_1 is given by

$$R_i^1 : \text{IF } x_1 \text{ is } A_i^1 \text{AND } x_2 \text{ is } A_i^2 \text{ THEN } z \text{ is } C_i \quad (1)$$

and the rule base of the fuzzy system S_2 is

$$R_i^2 : \text{IF } x_3 \text{ is } A_i^3 \text{AND } z \text{ is } C_i \text{ THEN } y \text{ is } B_i \quad (2)$$

The problem is that the properties and values of the output variable z are unknown. Thus neither the fuzzy clustering in the product space nor the fuzzy clustering in the input space with the output parameter fitting can be directly used. Subsystems at the first level have defuzzification procedure for all inputs, a rule base with a inference mechanism. The subsystem at the last level has a defuzzification block, a fuzzification block for inputs that are not input from subsystem at lower levels and a rule base with an inference mechanism. The internal structure of the simple hierarchical fuzzy systems is in Fig. 4. The fuzzy system implementation uses two sided Gaussian membership function. The two side Gaussian function is smooth, continuously

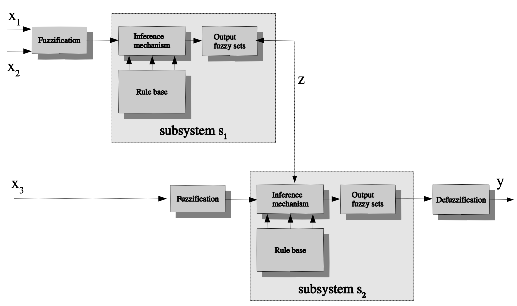

Fig. 4. The internal structure of the hierarchical fuzzy system

differentiable and the support of the Gaussian function is also infinitely large. The singleton model is used with the following inference mechanism

$$y(x) = \frac{\sum_{i=1}^{M} y_i \beta_i(x)}{\sum_{i=1}^{M} \beta_i(x)} \qquad (3)$$

where y_i is the position of i^{th} singleton, β_i is the firing strength of i^{th} rule.

The solution described below has been used to overcome mentioned problems. The presented approach is shown on the following example:

Example 1. Consider the non-linear function $y = x_1(1 - x_2) + x_3$. *Training set contains 1500 samples, the inputs* x_1, x_2 *and* x_3 *are randomly generated values* $x_i \in \langle 0, 1\rangle, i = 1, 2, 3$. *Testing set has also 1500 but in order to illustrate the method* $x_i, i = 1, 2, 3$ *are mutually independent sinus functions.*

Let analyze this nonlinear function from the fuzzy point of view. Consider that all inputs can reach two fuzzy terms - **small** and **large**. There are three fuzzy terms at the output - **small, medium** and **large**. The conventional fuzzy system proposed by Wang algorithm which cover the whole product space would have $2^3.3 = 24$ rules. The non-linear function can be decomposed into two parts.

$$y = x_1(1 - x_2) + x_3 \qquad (4)$$
$$y = z + x_3 \qquad z = x_1(1 - x_2)$$

It corresponds to the mathematical operation sequence, at first the multiplication then the addition. Analyzing the first part of the non-linear function $x_1(1 - x_2)$ the following rule base is to be most natural: Tab. 1 shows the fuzzy notation of the multiplication on universe $\langle 0, 1\rangle$. The subsystem S_2 represents addition $z + x_3$, its fuzzy notation is in Tab. 2. Rule bases in Tab. 1 and 2 are the fuzzy notation of the system in the example 1. Such a fuzzy system would probably choose a fuzzy expert to model the non-linear function and this is a reason why the structure

Tab. 1. Fuzzy system S_1 - rule base I

IF	x_1 is	AND	x_2 is	THEN	z is
R_1	small		small		small
R_2	small		large		small
R_3	large		small		large
R_4	large		large		small

of the hierarchical fuzzy system depicted in Fig. 3 has been chosen.

Tab. 2. Fuzzy system S_2 - rule base II

IF	x_3 is	AND	z is	THEN	y is
R_1	small		small		small
R_2	small		large		medium
R_3	large		small		medium
R_4	large		large		large

4. INTERCONNECTIONS

The output of the subsystem S_1 is used as the input thus is necessary to pass results to the next layer. Two method has been tested

- winner takes all
- weighted winner

"Winner takes all" is the method where the input pattern is passed through the subsystem S_1 then the membership to all output groups is measured. The number of the output groups with the highest membership is used as a result in the next layer. The inner variable z and the map S_1 have following meaning

design: (5)

$$S_1: \quad R^2 \longrightarrow Z \quad , \quad z = f(x_1, x_2)$$
$$z = \arg\max_k (A_k(x_i))$$

simulation: (6)

$$S_1: \quad R^2 \longrightarrow R^m \quad , \quad \mathbf{z} = f(x_1, x_2)$$
$$\mathbf{z} = [A_1(x_i), \ldots, A_m(x_i)]$$

where x_i is a particular datum, m is the number of clusters at the output of the subsystem S_1 and A_i are membership functions represented by multidimensional clusters. In our case $z \in (1, 2, 3, 4)$. This approach supposes that a clustering algorithm used in the next layer is able to deal with crisp value in this dimension because the groups can not be ordered. These four clusters represent some consequent terms (e.g. low, medium, big) but need not. At the second layer the pattern belongs to the cluster with the membership equal to one.

The second method is similar but the resulting cluster is passed with the weight which is equal to the membership $A_k(x_i)$ of the i^{th} data sample to the k^{th} cluster. The clustering in the second layer is processed in the same way, e.g. the crisp value representing the number of the cluster is used as an input. But real memberships fire corresponding rule propositions during the model simulation. While in the previous method only rules containing the rule proposition with the excited cluster are fired so in this method all rules can be fired.

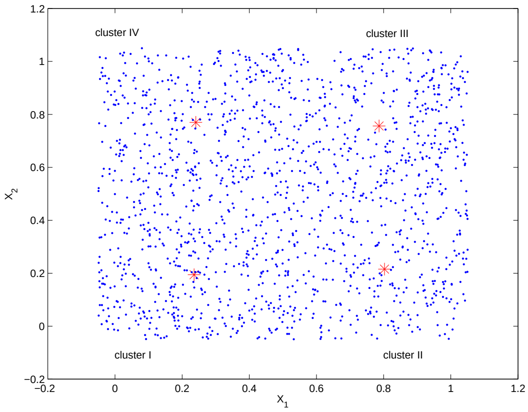

Fig. 5. Inputs of S_1 with centers of clusters (big asterisks)

The input space of the subsystem S_1 is divided to several groups using a fuzzy clustering algorithm. Any fuzzy clustering algorithm can be used at this stage (e.g. see (Bezdek, 1981; Höppner and Runkler, 1999)). Thus the variable z is represented by several clusters, in our case four clusters with the centers marked in Fig. 5. Because there is no knowledge on the variable z the clusters can not be lined up to create the standard universe of the variable z. The output of the subsystem S_1 is not a single value but at least two values - cluster and membership. The complete information is represented by memberships to all clusters, 4 membership in our example. This information is passed from the subsystem S_1 to the subsystem S_2. The basic structure of the rule base of the subsystem S_1 is in the Tab. 3. Data from the first

Tab. 3. Fuzzy rule base, subsystem S_1, basic structure - example

IF	x_1 is	AND	x_2 is	THEN	z is
R_1	small		small		cluster I
R_2	small		large		cluster III
R_3	large		small		cluster II
R_4	large		large		cluster IV

Tab. 4. Fuzzy rule base, subsystem S_2, basic structure - example

IF	z is	AND	x_3 is	THEN	y is
R_1	cluster I		large		medium
R_2	cluster II		large		large
R_3	cluster III		small		small
R_4	cluster II		small		medium
R_5	cluster IV		large		medium
R_6	cluster I		small		small
R_7	cluster III		large		medium
R_8	cluster IV		small		small

layer are passed to the second layer and a new data set is clustered and the rule base II of the subsystem S_2 is established. This structure is summarized in Tab. 4. Now the proposed hierarchical model has four rules in the subsystem S_1 at the first layer and eight rules in the subsystem S_2 at the second layer. Both rule bases could be analyzed and simplified. The next section addresses this problem.

5. RULE BASE SIMPLIFICATION

As mentioned in the previous sections, the meaning of clusters at the connection z is unknown, clusters could not be denoted by fuzzy terms and ordered. But we can presume that clusters have some meaning, e.g. clusters I and II mean small and clusters III and IV mean large. Merging clusters in the rule base would simplify the structure but the consistency of rules must be retained. By merging the cluster I with the cluster II the inconsistent rule base arises because rules R_1 and R_2 have the same premise " **IF** z IS new cluster **AND** x_3 IS large" but different conclusions (R_1 - medium, R_2 - large). If the cluster I is joined with the cluster III and the cluster IV (new one is denoted as cluster V) then the rule base has only four rule given by Tab. 5. Comparing tables Tab. 1 to 3 and 2 to Tab. 5 it

Tab. 5. Fuzzy rule base, subsystem S_2, after simplification - example

IF	x_1 is	AND	x_2 is	THEN	y is
R_1	cluster V		large		medium
R_2	cluster II		large		large
R_3	cluster V		small		small
R_4	cluster II		small		medium

is clearly seen that the cluster II represents fuzzy term large while the cluster V (it is clusters I,III and IV) means small. In such a case the assumed fuzzy rules are the same as rules obtained by fuzzy clustering method from data.

The graphical results of this algorithm with the "weighted winner" connection between subsystems are depicted on Fig. 6. The resulting hierarchical structure has eight rules (4 at the first stage + 4 at the second stage)

The rule base simplification is the reason why the topology of the fuzzy hierarchical system is restricted to tree. In the case of branching of the variable z_1 in Fig. **??** there could be problems with meanings of clusters. Assume that there are four clusters at the output of the rule base I. After the simplification of the rule bases III and IV two possible interpretation of clusters would be established. For instance in the case of the rule base III three of four mean **small** and the remaining means **large** while in the case of the rule base IV one half mean **large**, the rest mean **small**. Both specific rule bases are consistent but the whole hierarchical fuzzy system is not.

Algorithm - model design

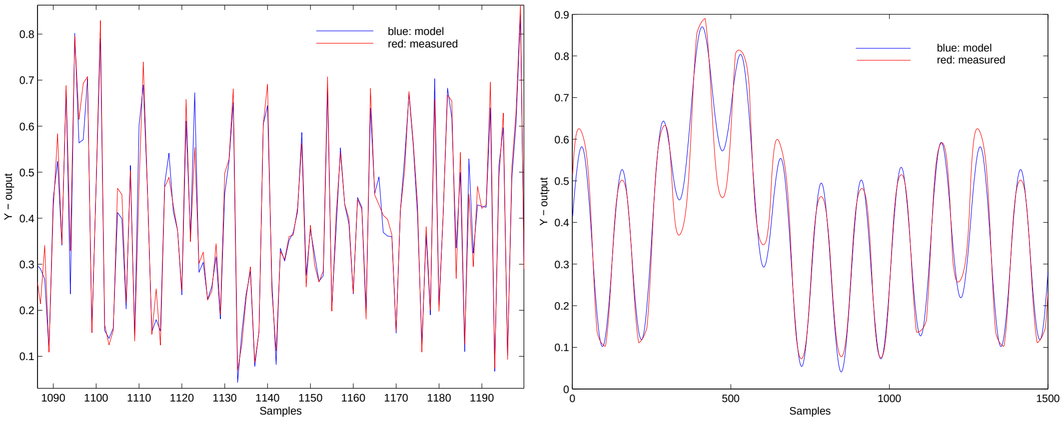

Fig. 6. The example with "weighted winner" approach, training data set - left (zoom) and test data set - right

Set the structure of the fuzzy hierarchical system, S_{ij} is a j^{th} subsystem in i^{th} layer. The structure has m layers and n_i subsystems at i^{th} layer. For the sake of simplicity MISO system is supposed. Let denote inputs x and outputs y, U_{ij} is the input matrix, it can contain inputs x_i or inner variables z_{ij} which is the output of the subsystem S_{ij}.
An input membership function A_i^j is i^{th} membership function of j^{th} input, B_i are output singletons, C_{ij}^k is k^{th} cluter in ith inner layer at jth output.

for i = 1 : $m-1$

for j = 1: $n_i - 1$

Generate clusters at the i^{th} level using the matrix U_{ij}

Create the rule base of the fuzzy subsystem S_{ij}

Generate z_{ij} according to

$$z_{ij}(x) = \arg\max_k (C_{ij}^k(x)) \qquad (7)$$

end

Prepare data inputs for the next level regarding to the given structure. Symbolic description

$$\mathbf{U}_{i+1} = [x, z_i] \qquad (8)$$

end

Set $i = m$

Use any clustering algorithms to generate subsystems at the highest level

Adjust S_{m1}

Algorithm - model simulation

Suppose the fuzzy hierarchical system built according to the previous algorithm.

for i = 1 : $m-1$

for j = 1: $n_i - 1$

If there are input from lower levels then particular propositions of fuzzy rules are excited by

$$R_2^i : \text{IF } x_3 \text{ is } A_1^i \text{AND } \underbrace{z \text{ is } C^i}_{C_{i-1}(x)}$$

$$\text{THEN } y \text{ is } B^i \qquad (9)$$

Compute $\mathbf{z}_{ij}$

$$\mathbf{z}_{ij} = [C_1(x), \ldots, C_k(x)] \qquad (10)$$

where k is the number of the output clusters.

end

Prepare inputs for the next level regarding to the given structure. Symbolic description

$$\mathbf{U}_{i+1} = [x, \mathbf{z}_i] \qquad (11)$$

end

Set $i = m$

Use inference mechanism of the subsystem at the last level to compute the system output.

6. EXAMPLES

This section presents the use of the described method in the prediction of the gas consumption. The goal of the fuzzy model is to predict the daily gas consumption for a year in the Czech Republic for different weather scenarios (hot summer, middle summer, very cold winter etc.). The more detailed description of this project can be found in (Šindelář and Pavlík, 2003).

In this project two groups of inputs are available. The first group consist of measured weather information as the normal temperature and the average temperature. The calendar information are in the second group. Such information as day of a week (Monday, ..., Sunday), holidays, days and months etc. can be found here. Using the input selection algorithm proposed in (Šindelář and Vlček, 19-21 September 2002), the following inputs have been chosen

- temperatures in day k, $k-2$, $k+2$
- day of a week
- temperature normal in day k

The attribute to be predicted in terms of above six inputs is the daily gas consumption in the Czech Republic. Available data sets are divide into two sets -

Tab. 6. Fuzzy models comparison

fuzzy model	number of rules	MAPE [%]	zone error class I	class II	class III
conventional FM	60	6.41	47.1	30.4	22.5
hierarchical FM I	5 + 15	6.89	46.2	30.8	23.0
hierarchical FM II	18 + 35	6.38	53.9	24.5	21.6

training set and testing set. The training set is used to develop a fuzzy model while the testing set is used to verify the quality of the fuzzy model.

To compare the model quality two criteria are applied. The mean average percentage error is defined by

$$MAPE = \frac{1}{N}\sum_{i}^{N}\frac{|y_i - \hat{y_i}|}{y_i}100\% \qquad (12)$$

where y_i is the measured output and $\hat{y_i}$ is the model output. The second criteria is the zone error which counts the sample belonging to classes defined by

(1) class I- $\frac{|x_i - \hat{x_i}|}{x_i}100 \leq 5$ %

(2) class II - 5 % $< \frac{|x_i - \hat{x_i}|}{x_i}100 \leq 10$ %

(3) class III - $\frac{|x_i - \hat{x_i}|}{x_i}100 > 10$ %

Developing the conventional fuzzy model lead to results shown in the first row of Tab. 6. The internal structure of the hierarchical system is given by the nature of inputs, see Fig 7. The measured temperatures are inputs to the subsystem S_1 at the first stage. The output of this subsystem z together with the calendar information are inputs to the subsystem S_2. In the rows 2 and 3 of Tab. 6 the results of two different hierarchical models are shown. The fuzzy model I has 20 rules in total, the subsystem S_1 contains 5 rules and the subsystem S_2 contains 15 rules. The numerical quality is comparable with the conventional fuzzy model but the total number of the rule is three times lower. The hierarchical fuzzy system II has the lower number of rules as the conventional system but 8% more samples have MAPE lower than 5% which the most favourable in this case because the errors over 5% are penalized.

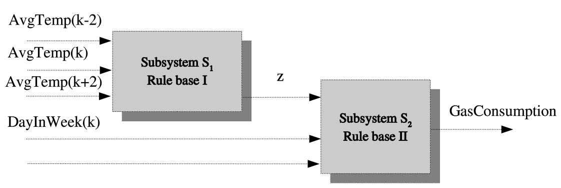

Fig. 7. Internal structure of fuzzy model for gas consumption prediction

7. CONCLUSIONS

The very simple method for fuzzy hierarchical systems design is proposed in this paper. Most of methods start from the analyze of the conventional fuzzy model. This approach requires developing the initial fuzzy model that is usually large and very complex. The presented method overcomes this problem by the sequential generating layer components of hierarchical fuzzy system from data. The most problematic part of the hierarchical fuzzy system are connections between layers because data sets are not available for these variables. To get the data set for fuzzy clustering algorithm at higher levels, inputs entering to lower level are clustered in the input space using an clustering algorithm,it means the data are divided in the desired number of groups. Each data sample is assigned to some of these groups and this information is passed to next levels. Interconnections during the model design are the weak point of the method. The future research should be focused on a better coding of inner variables.

ACKNOWLEDGMENT

This work was in part supported by the Ministry of Education of the Czech Republic under Project LN00B096.

REFERENCES

Bezdek, J.C. (1981). *Pattren Recognition with Fuzzy Objective Function Algorithms*. Plenum. New York.

Gegov, A. E. and P. M. Frank (1995). Decomposition of multivariable systems for distributed fuzzy control. *Fuzzy Sets and Systems* pp. 329 – 340.

Höppner, F. and T. Runkler (1999). *Fuzzy cluster analysis*. John Wiley. New York.

Joo, Moon G. (2003). A method of converting conventional fuzzy logic system to 2 layered hierarchical fuzzy system. *IEEE International Conference on Fuzzy Systems Proceedings 2003* pp. 1357–1362.

Šindelář, R. and P. Pavlík (2003). Fuzzy models for gas distribution network optimization. *WSEAS Transaction on Curcuits and Systems* **2**, 692–697.

Šindelář, R. and Z. Vlček (19-21 September 2002). Input selection for fuzzy models. *The European Symposium on Intelligent Technologies, Hybrid Systems and their implementation on Smart Adaptive Systems, Albufeira, Portugal*.

Wang, L.-X. (1998). Universal approximation by hierarchical fuzzy systems. *Fuzzy Sets and Systems* **2**, 223–230.

Wang, L. X. (1999). Analysis and design of hierarchical fuzzy systems. *IEEE Transaction of Fuzzy Systems* **7**, 617 – 624.