Design of Transparent Mamdani Fuzzy Inference Systems

Giovanna CASTELLANO, Anna M. FANELLI and Corrado MENCAR
CILAB – Computational Intelligence LABoratory
Department of Computer Science – University of Bari – Bari, ITALY
e-mail: {castellano, fanelli, mencar}@di.uniba.it

Abstract: In this paper, we propose a technique to design Fuzzy Inference Systems (FIS) of Mamdani type with transparency constraints. The technique is based on our Crisp Double Clustering algorithm, which is able to discover transparent fuzzy relations that can be directly translated into a human understandable rule base. As a key feature, the user can tune the granularity level of the rule base so as to properly balance the trade off between accuracy and transparency. The resulting FIS bears a transparent knowledge base that can be easily understood by human users and can be effectively used to solve soft computing problems. The work is accompanied by an illustrative example that show the validity of the approach.

1 Introduction

Fuzzy Inference Systems (FIS) are popular computing frameworks based on the concepts of fuzzy set theory, which have been applied with success in many fields like control, decision support, system identification, etc. [13]. Their success is mainly due to their closeness to human perception and reasoning, as well as their intuitive handling and simplicity, which are important factors for acceptance and usability of the systems [12].

In fig. 1 the general schema of a FIS is portrayed. In particular, three main modules are of particular interest: a fuzzifier, a rule base and a defuzzifier. While the fuzzifier and the defuzzifier have the role of converting external information in fuzzy quantities and vice

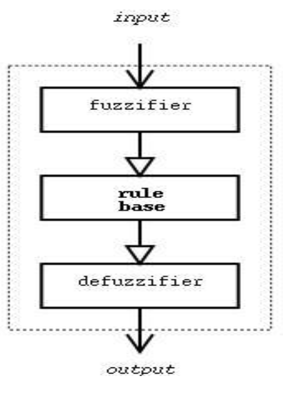


Figure 1: *Scheme of a Fuzzy Inference System*

versa, the core of a FIS is its knowledge base, which is expressed in terms of fuzzy rules and allows for approximate reasoning [5].

Typically, a FIS can be classified according to three types of models that are distinguished in the formalization of the fuzzy rules [7]. In this work, we focus on the Mamdani type, which is characterized by the following fuzzy rule schema:

$$\text{IF } \boldsymbol{x} \text{ is } \boldsymbol{A} \text{ then } \boldsymbol{y} \text{ is } \boldsymbol{B} \tag{1}$$

where $\boldsymbol{A}$ and $\boldsymbol{B}$ are fuzzy sets defined on the input and output domains respectively.

Mamdani FIS type was proposed as the first attempt to solve control problems by a set of linguistic rules obtained from experienced human operators [9]. The main feature of such type of FIS is that both the antecedents and the consequents of the rules are expressed as linguistic constraints [17]. As a consequence, a Mamdani FIS can provide a highly intuitive knowledge base that is easy to understand and maintain, though its rule formalization requires a time consuming defuzzification procedure. For such reasons, Mamdani type FISs can be used as valid supports in all such fields – like Medicine, Economics, etc. – where transparency more than cutting-edge efficiency is required.

In many real-world problems, the input/output mapping must be acquired from data. In such case, the design of Mamdani FISs must take into account both accuracy and transparency requirements in order to avoid the acquisition of a knowledge base inaccessible by human users. Moreover, the design procedure should involve the user in choosing the trade off level between accuracy and transparency. Such choice should be based on selecting appropriate values of hyper-parameters with immediate meaning.

While the accuracy issue has been tackled as soon as FISs were introduced, the transparency requirement was analyzed only recently by several authors with different approaches [2][6][8][10][12][14][15][16]. In this paper, we propose a technique to automatically design Mamdani FISs from available data that also satisfy the transparency requirement. Such requirement is formalized as a set of formal constraints commonly adopted in literature (see e.g. [16]), which guide the learning process for acquiring a rule base from data. As a key feature, the user can decide the maximum number of rules, as well as the proper level of granularity of each input, so as to calibrate the transparency/accuracy trade off in a simple yet effective way.

The design technique involves a learning strategy based on our Crisp Double Clustering (CDC) algorithm [4], [3], which is able to automatically discover transparent fuzzy clusters from data at various degrees of granularity, according to user choices. The learning strategy is defined by three main steps. First, a Vector Quantization scheme is applied on the available data in order to provide a compressed yet informative representation of data in the input/output product space. As a second step, the derived multidimensional prototypes are projected onto each axis and further clustered so as to join similar projections in the same cluster. As a final step, one-dimensional clusters are combined to form linguistically interpretable fuzzy sets in the input/output product space. Such multidimensional fuzzy sets are directly translated into fuzzy rules of Mamdani type, which result both accurate and transparent.

The paper is organized as follows. In Section 2, the structure of Mamdani FIS type is described in deeper detail and transparency constraints are introduced. In Section 3, a description of the design technique is provided, while in Section 4 some simulation results are shown to validate the proposed approach. Finally, in Section 5 some conclusive remarks are drawn.

2 Transparent Mamdani FIS

From a functional point of view, a Mamdani FIS is a nonlinear mapping from an input domain $\mathbf{X} \in \mathbb{R}^n$ to an output domain $\mathbf{Y} \in \mathbb{R}^m$. Without loss of generality, we assume that both the input domain and the output domain are defined as hyper-intervals:

$$\mathbf{X} \times \mathbf{Y} = \mathbf{Z} = Z_1 \times \cdots \times Z_n \times Z_{n+1} \times \cdots \times Z_{n+m} \tag{2}$$

where $Z_i = [m_i, M_i]$, $i = 1, 2, ..., n+m$.

This input/output mapping is realized by means of R rules of the following form:

$$\text{IF } \mathbf{x} \text{ is } \mathbf{A}^{(r)} \text{ THEN } \mathbf{y} \text{ is } \mathbf{B}^{(r)} \tag{3}$$

where $r = 1, 2, \ldots, R$ is the index of the rule, while $\mathbf{A}^{(r)}$ and $\mathbf{B}^{(r)}$ are fuzzy relations over $\mathbf{X}$ and $\mathbf{Y}$ respectively. When an input vector $\mathbf{x}$ is presented to the system, a fuzzy set $\mathbf{B}$ is inferred according to the following relation:

$$\mathbf{B}(\mathbf{y}) = \bigvee_{r=1}^{R} \left(\mathbf{A}^{(r)}(\mathbf{x}) \wedge \mathbf{B}^{(r)}(\mathbf{y}) \right) \tag{4}$$

where the the formalism $A(\cdot)$ denotes the membership function of a fuzzy set A, and $\wedge, \vee$ are a T-norm and a T-conorm respectively (usually the *min* and the *max* functions are used).

When a crisp output is required, the resulting fuzzy set can be defuzzified by means of several strategies, among which the most frequently used is the centroid of area, defined as:

$$\tilde{\mathbf{y}} = \frac{\int_{\mathbf{Y}} \mathbf{B}(\mathbf{y}) \cdot \mathbf{y} \cdot \mathbf{dy}}{\int_{\mathbf{Y}} \mathbf{B}(\mathbf{y}) \cdot \mathbf{dy}} \tag{5}$$

Variants of Mamdani type can be attained with different forms of defuzzification, as well as different realizations of the T-norm and the T-conorm. For further details, the reader is referred to [7].

The transparency property of fuzzy rules is directly related to the involved fuzzy sets. In particular, a set of formal constraints must be satisfied in order to achieve such requirement. A generally adopted constraint requires that each fuzzy relation is expressed as a Cartesian product of one-dimensional fuzzy sets, so as to represent a conjunction of qualities defined on each variable. Stated formally, this constraint can be expressed as follows:

$$\forall r = 1, 2, \ldots, R : \begin{matrix} \exists A_1^{(r)}, \ldots, A_n^{(r)} \\ \exists B_1^{(r)}, \ldots, B_m^{(r)} \end{matrix} \; s.t. \; \begin{matrix} \mathbf{A}^{(r)} = A_1^{(r)} \times \cdots \times A_n^{(r)} \\ \mathbf{B}^{(r)} = B_1^{(r)} \times \cdots \times B_m^{(r)} \end{matrix} \tag{6}$$

In addition, the involved one-dimensional fuzzy sets must satisfy a set of constraints so as to associate a clear semantics that is expressed by meaningful linguistic labels. Specifically, for each family Φ_i, $i = 1, 2, \ldots, n+m$ of fuzzy sets defined on the same domain Z_i (input or output), in this work we adopt the following constraints:

- Each fuzzy set must be normal, unimodal and convex, so that each fuzzy set is fully represented by only one prototype and all the other elements have a membership grade that decreases as their similarity with the prototype decreases:

$$\begin{gathered} \forall A \in \Phi_i \exists | x \in Z_i \; s.t. \; A(x) = 1, \\ \forall A \in \Phi_i \forall x, y, z \in Z_i : x < y < z \rightarrow A(y) \geqslant min\{A(x), A(z)\} \end{gathered} \tag{7,8}$$

- Fuzzy sets should cover the entire domain so that each data is well represented by some linguistic terms. Usually, a strict coverage is preferable, which requires that for each element of the domain there exist at least one fuzzy set whose membership grade is greater than a predefined threshold ϵ :

$$\forall x \in Z_i \exists A \in \Phi_i \, s.t. \, A(x) \geq \epsilon \tag{9}$$

- Two fuzzy sets must not overlap too much, so that they have distinct semantic meanings. Such property can be formalized by the possibility function, which should be smaller than a predefined threshold for each couple of fuzzy sets. In this work, the possibility threshold coincides with the coverage threshold:

$$\forall A, B \in \Phi_i : \sup_{x \in Z_i} \min\{A(x), B(x)\} \leq \epsilon \tag{10}$$

- The extreme values of each domain must be prototypes for some fuzzy sets, which are called leftmost and rightmost fuzzy sets respectively:

$$\exists A, B \in \Phi_i \, s.t. \, A(m_i) = 1 \wedge B(M_i) = 1 \tag{11}$$

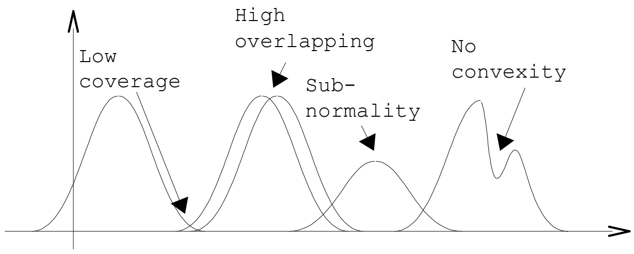


Figure 2: *An example of fuzzy sets violating some transparency constraints*

In fig. 2, an example of fuzzy sets violating some transparency constraints is illustrated. For such fuzzy sets, the association of meaningful linguistic labels is difficult or counterintuitive. On the contrary, fuzzy sets in fig. 3 satisfy all transparency constraints and linguistic labels can be associated immediately.

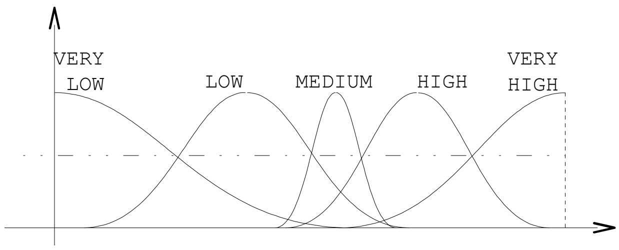


Figure 3: *An example of fuzzy sets satisfying all transparency constraints*

The above described transparency constraints are further integrated by some guidelines that suggest a low number of fuzzy sets per input and a low number of rules. However, the lower are such hyper-parameters, the lower is the overall accuracy. As a consequence, the choice of these values should be ascribed to the user, who may decide toward a simple model that roughly describes the system behavior, rather than a system with a more accurate, though less immediate, knowledge base.

3 The design technique

The design of a Mamdani FIS with transparency property must take into account the above defined transparency constraints. If Gaussian membership functions are chosen as a common shape for all fuzzy sets, then normality, unimodality and convexity constraints are always guaranteed. Gaussian membership functions are fully determined by two parameters, namely the center ω and the width σ, and have the following functional form:

$$A(x)=\exp\left(-\frac{(x-\omega)^2}{2\sigma^2}\right) \tag{12}$$

The remaining transparency constraints can be guaranteed only if a proper technique for the construction of fuzzy relations is adopted. A common solution is to apply a clustering technique on the product space $\mathbf{Z}$ (see e.g. [1], [11]). However, if the clustering algorithm is not properly chosen, interpretability of the resulting fuzzy relations can be hampered due to the violation of transparency constraints.

In the following, we describe the proposed technique to automatically design Mamdani FISs from available data that also satisfy the transparency requirement. Such technique is based on our CDC algorithm, which is particularly suited for generating multidimensional fuzzy relations that satisfy the aforementioned interpretability constraints. The CDC-based learning strategy is defined by three main steps:

1. First, a vector quantization algorithm is applied so as to obtain a number of multidimensional prototypes $\mathbf{P}=\{\mathbf{p}_1,\mathbf{p}_2,\ldots,\mathbf{p}_K\}\subset\mathbf{Z}$ that represent data in a compressed yet informative way. The number K of prototypes corresponds to the maximum number of rules and is provided by the user, who can therefore prefer a more accurate vs. a more readable knowledge base.
2. As a second step, the multidimensional prototypes are projected on each dimension, where they are clustered into a number of one-dimensional clusters per axis $C_i=\{c_{i1},c_{i2},\ldots,c_{i,K_i}\}\subset[m_i,M_i]$, $i=1,2,\ldots,n+m$. The number of clusters is selected by the user, who can therefore choose a proper description level of data. Midpoints between adjacent clusters $t_{ij}=(c_{i,j}+c_{i,j+1})/2$ define intersection points between adjacent membership functions and are retained for the subsequent step;
3. Finally, one-dimensional Gaussian fuzzy sets are first defined onto each axis so that interpretability constraints are satisfied. The centers and widths of each fuzzy sets defined for the *i*-th dimension are defined as follows:
   $$\omega_{ij}=(t_{ij}+t_{i,j+1})/2 \tag{13}$$
   $$\sigma_{ij}=(t_{i,j+1}-t_{ij})/(2\sqrt{-\ln\epsilon}) \tag{14}$$
   Fuzzy relations are formed as Cartesian product of one-dimensional fuzzy sets and expressed as conjunction of linguistic labels. To avoid combinatorial explosion, only those relations that represent multidimensional prototypes are retained, while all the others are discarded.

The result of the learning algorithm is a family of fuzzy relations defined on the product space $\mathbf{Z}$. For each fuzzy relation $\mathbf{R}^{(r)}=A_1^{(r)}\times\cdots\times A_n^{(r)}\times B_1^{(r)}\times\cdots\times B_m^{(r)}$ a fuzzy rule is straightforwardly derived based on the following schema:

$$\begin{array}{l}\text{IF } x_1 \text{ is } A_1^{(r)} \text{ AND} \cdots \text{AND } x_n \text{ is } A_n^{(r)} \\ \text{THEN } y_1 \text{ is } B_1^{(r)} \text{ AND} \cdots \text{AND } y_m \text{ is } B_m^{(r)}\end{array} \tag{15}$$

Since input and output fuzzy sets are derived from the same Cartesian product and successively split in the antecedent part and the consequent part, it is preferable that the T-norm of the inference rule (4) coincides with the conjunction operator used to evaluate the antecedent truth value.

The proposed clustering algorithm generates transparent fuzzy models, since all identified fuzzy sets satisfy the transparency constraints defined in the previous section, so the derived knowledge base is interpretable by human users.

4 Illustrative example

As an example to show the effectiveness of the proposed design technique, we consider the problem of predicting future values of the time series defined by the following Mackey-Glass differential equation:

$$\frac{dx}{dt}(t)=\frac{0.2x(t-17)}{1+x(t-17)^{10}}-0.1x(t) \tag{16}$$

In particular, the problem requires the prediction of $x(t+6)$, when $x(t-18)$, $x(t-12)$, $x(t-6)$ and $x(t)$ are given. We generate a dataset of 1000 samples for $t=118,\ldots,1117$, with the initial condition $x(0)=1.2$ and then we split the dataset in a training set (first 500 samples) and a test set (last 500 samples). Moreover we fix the coverage threshold to 0.5 and we adopt product and max functions as T-norm and T-conorm respectively.

By varying the number of multidimensional prototypes and fuzzy sets per input, different levels of accuracy are achieved. In particular, with 3 fuzzy sets per input and 3 multidimensional prototypes, the design procedure carries out a FIS with only 3 rules and a rough RMSE of 0.1281 on the test set. Despite the high error rate, this simple FIS provides an approximate outline of the system behavior. In table 1 the derived knowledge base is portrayed, while in fig. 4 membership functions of the involved fuzzy sets are illustrated.

Table 1: *Derived knowledge base for the 3-rules FIS*

IF x(t-18) is LOW and x(t-12) is LOW and x(t-6) is LOW and x(t) is MED[1] THEN x(t+6) is HIGH
IF x(t-18) is MED and x(t-12) is MED and x(t-6) is HIGH and x(t) is HIGH THEN x(t+6) is MED
IF x(t-18) is HIGH and x(t-12) is HIGH and x(t-6) is MED and x(t) is LOW THEN x(t+6) is LOW

By increasing the number of fuzzy sets per input and the number of multidimensional prototypes, the accuracy of the resulting FIS drastically improves, but with the drawback of a sensible increase in the number of rules that hampers the legibility of the discovered knowledge base. As an example, with 7 fuzzy sets per input and 200 prototypes, the resulting FIS embodies a knowledge base of 127 fuzzy rules and provides a RMSE of 0.0342 on the test set. In fig. 5, the approximation results of the 3-rules FIS and the 127-rules FIS are displayed. As expected, while the 3-rules FIS provides only information on the trend of the time series, the 127-rules FIS attains an almost perfect fit. The RMSE of the 127-rules FIS is comparable with results obtained by NEFPROX on the same simulation environment [11] and the same number of rules. However, our results are automatically obtained by simply choosing the maximum number of rules and the number of fuzzy sets per input, without forcing the user to select appropriate parameters to obtain the desired results.

[1]MED stands for MEDIUM

5 Conclusions

We have presented a technique for automatically design FISs of Mamdani type by satisfying a set of formal constraints commonly adopted to define fuzzy sets transparency. As a key feature of the proposed approach, the user can decide the proper trade-off between accuracy and transparency by providing the maximum number of desired rules, as well as the number of fuzzy sets per each input. Such feature can be advantageously used in a multi-level strategy, where a highly transparent FIS with few rules is used for an overall description of the system behavior, while a more accurate FIS is employed for effective system modeling.

The proposed technique reveals a promising research direction toward bridging the gap between the numerical world, in which often data are measured and gathered, and the symbolic world used to represent transparent knowledge. As a preliminary approach, the technique can be further refined by automatically maximizing both accuracy and transparency. Solutions in such direction are under examination.

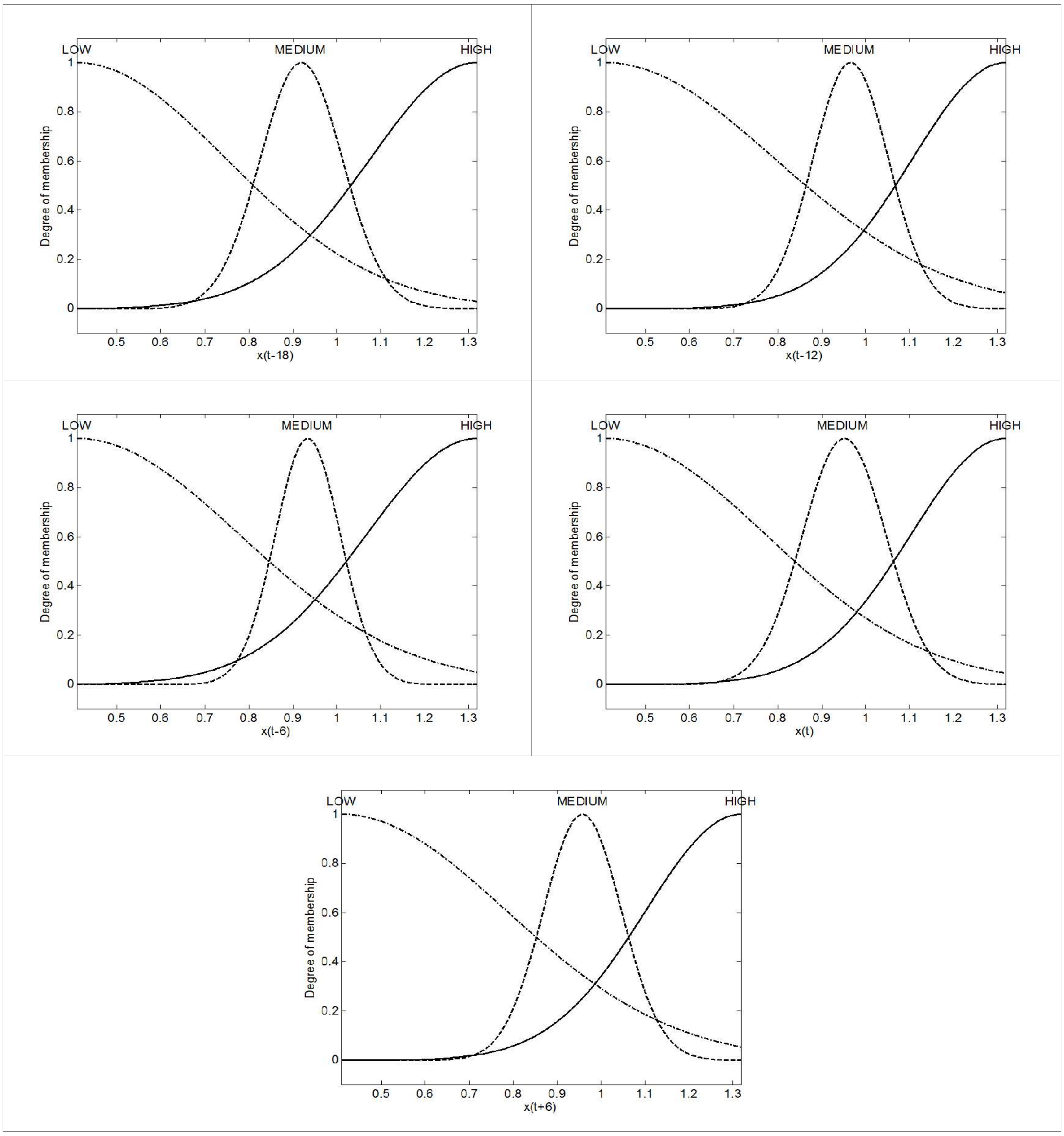
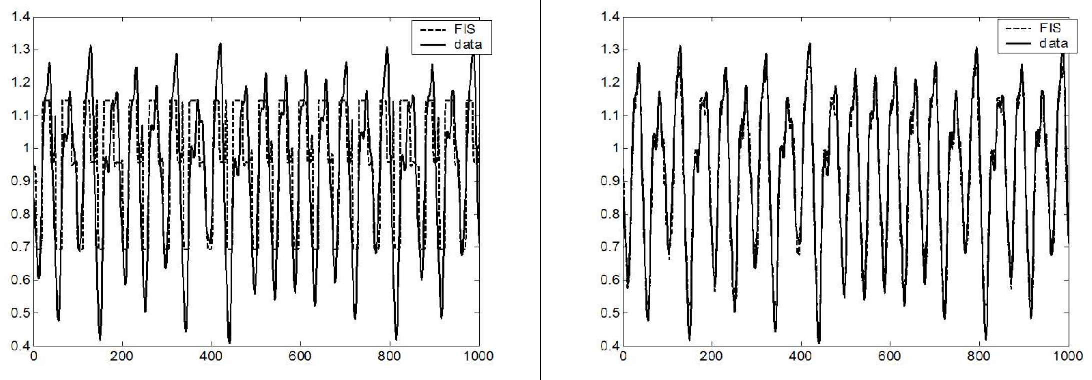


Figure 4: *Fuzzy sets defined for the 3-rules Mamdani FIS*



(a) (b)

Figure 5: *Approximation of the Mackey-Glass time series by the 3-rules FIS (a) and the 127-rules FIS (b)*

6 References

[1] Babuska, R.; Verbruggen, H.B., "Constructing fuzzy models by product space clustering", in Hellendoorn, H. and Driankov, D. (eds.), *Fuzzy Model Identification: Selected Approaches*, pp. 53-90, Springer, Berlin, Germany, 1997

[2] Casillas, J.; Cordòn, O.; Herrera, F.; Magdalena, L. (eds.), *Interpretability Issues in Fuzzy Modeling*, Springer, 2003

[3] Castellano G.; Fanelli, A.M.; Mencar, C., "Fuzzy Granulation of multi-dimensional data by a Crisp Double-Clustering algorithm*", in Proc. of 7th Multi-Conference on Systemics, Cybernetics and Informatics (SCI2003)*, Orlando, Fl, US, July 2003, pp.

[4] Castellano, G.; Fanelli, A.M.; Mencar, C., "Generation of interpretable fuzzy granules by a double-clustering technique", *Archives of Control Sciences: Special issue on Granular Computing*, vol. 12(4), pp. 397-410, 2002

[5] Czogala, E.; Leski, J., *Fuzzy and Neuro-Fuzzy Intelligent Systems*, Physica-Verlag, 2000

[6] Guillaume, S., "Designing Fuzzy Inference Systems from Data: An Interpretability Oriented Review", *IEEE Transactions on Fuzzy Systems*, vol. 9(3), pp. 426-443, IEEE, 2001

[7] Jang, J-S.R.; Sun, C-T., "Neuro-Fuzzy Modeling and Control", *Proceedings of IEEE*, vol. 83, pp. 378-406, IEEE, 1995

[8] Jin, Y.; von Seelen, W.; Sendhoff, B., "On Generating FC^3 Fuzzy Rule Systems from Data Using Evolution Strategies", *IEEE Transactions on Systems, Man and Cybernetics, part B*, vol. 29(6), pp. 829-845, IEEE, 1999

[9] Mamdami, E.H.; Assilina, S., "An experiment in linguistic synthesis with a fuzzy logic controller", *International Journal of Man-Machine Studies*, vol. 7(1), pp. 1-13, 1975

[10] Marìn-Blàzquez, J.G.; Shen, Q., "From Approximative to Descriptive Fuzzy Classifiers", *IEEE Transactions on Fuzzy Systems*, vol. 10(4), pp. 484-497, IEEE, 2002

[11] Nauck, D.; Kruse, R., "A Neuro-Fuzzy Approach to obtain Interpretable Fuzzy Systems for Function Approximation*", in Proc. of IEEE Int. Conf. on Fuzzy Systems (FUZZ-IEEE'98)*, Anchorage, AK, May 1998, pp. 1106-1111

[12] Nauck, D.; Kruse, R., "Obtaining Interpretable Fuzzy Classification Rules from Medical Data", *Artificial Intelligence in Medicine*, vol. 16, pp. 149-169, Elsevier, 1999

[13] Ross, T.J., *Fuzzy Logic with Engineering Applications*, McGraw Hill Int. Ed., 1997

[14] Roubos, H.; Setnes, M., "Compact and Transparent Fuzzy Models and Classifiers through Iterative Complexity Reduction", *IEEE Transactions on Fuzzy Systems*, vol. 9(4), pp. 516-524, IEEE, 2001

[15] Setnes, M.; Babuska, R.; Kaymak, U.; van Nauta Lemke, H.R., "Similarity Measure in Fuzzy Rule Base Simplification", *IEEE Transactions on Systems, Man and Cybernetics, part B*, vol. 28(3), pp. 376-386, IEEE, 1998

[16] Valente de Oliveira, J., "Semantic Constraints for Membership Function Approximation", *IEEE Transactions on Systems, Man and Cybernetics, part A*, vol. 29(1), pp. 128-138, IEEE, 1999

[17] Zadeh, L.A., "Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic", *Fuzzy Sets and Systems*, vol. 90, pp. 117-117, Elsevier, 1997