Modeling and Linguistic Knowledge Extraction of Systems using Fuzzy Relational Models

Fuzzy Sets and Systems

Vol. 121, pp. 113-126, 2001

Unformatted Manuscript
Modeling and Linguistic Knowledge Extraction from Systems using Fuzzy Relational Models

R. J. G. B. Campello and W. C. Amaral
DCA/FEEC/Unicamp
CP 6101, CEP 13083-970, Campinas, SP, Brazil
e-mail: (ricardo,wagner)@dca.fee.unicamp.br

Abstract Fuzzy relational models have been widely investigated and found to be an efficient tool for the identification of complex systems. However, little attention has been given to the linguistic interpretation of these models. The use of relational models is recommended since their development follows a natural sequence based on the original ideas about fuzzy sets and fuzzy logic, involving the estimation of the relations existing between linguistic terms which have previously been defined by the user. In the present paper the problem of extracting linguistic knowledge from systems by using relational models is addressed. A new algorithm for the identification of these models which can provide analytical or numerical solutions depending on user requirements is also proposed. Examples are presented showing that both quantitative and qualitative modeling can be effectively achieved by combining the proposed methodologies for identification and extraction of linguistic knowledge from systems.

Keywords: Fuzzy modeling, knowledge extraction, relational models, linguistic models.

1 Introduction

Since Zadeh introduced the concept of fuzzy behavioral algorithms for the approximate description of the behavior of systems using fuzzy conditional statements [28], fuzzy modeling has become a powerful tool for complex system identification when conventional methods do not provide an adequate performance. The most important characteristic of models based on fuzzy logic techniques (fuzzy models) is the ability to deal with both quantitative and qualitative information. This ability, for example, allows the development and enhancement of models using qualitative information provided by experts, such as process operators. Nevertheless, the most relevant aspect of this ability is that it makes it possible to extract linguistic knowledge from systems using data.

Following the basic structure of Zadeh’s fuzzy behavioral algorithms, Tong [25] used a fuzzy model for the identification of systems from numerical data. The kind of model used by Tong is currently called the linguistic model and
is constituted by a set of fuzzy rules of the format “If \( x \) is \( A \) THEN \( y \) is \( B \)”, where \( x \) and \( y \) are input and output variables, respectively, and \( A \) and \( B \) are linguistic terms assigned to fuzzy sets. This type of model is linguistically interpretable, since the rules are semantically clear. However, the problem of getting an adequate rule set from data is not a trivial task, since the rules involve linguistic terms (e.g., \([22, 21]\)).

An alternative structure for fuzzy models was suggested by Takagi and Sugeno [23]. The Takagi-Sugeno fuzzy model is constituted by a set of fuzzy rules whose consequents are (nonfuzzy) functions which map the model inputs onto the output. The parameters of these functions can be estimated using identification algorithms like the Least Squares (LS) methods or the Kalman Filter [10]. The main problem with this approach is that the models cannot provide linguistic knowledge about the systems to be identified, because the interpretability of the rules is lost. Setnes et al. [20], however, proposed a methodology for linguistic knowledge extraction from special (but limited) types of Takagi-Sugeno models in which the consequents of the rules are constants instead of linear or non-linear functions of the input variables. Two important considerations should be kept in mind concerning the work referred to above: The first is that, as highlighted by the authors, the special types of Takagi-Sugeno models mentioned are also specific cases of linguistic models in which the consequents are singletons. The second is that their methodology involves a sequence of steps with estimation of the constant consequents of the rules (using LS), followed by grouping of these consequents according to some criterion and, finally, interpretation of the resulting groups, i.e., assignment of linguistic terms to them. However, based on the original ideas about fuzzy sets and fuzzy logic [28, 29], it is natural to assume that a more adequate sequence for linguistic knowledge extraction is the \( a \ priori \) definition of the linguistic terms (fuzzy sets) of the input and output variables, followed by the estimation of the relations (rules) existing between these terms, as will be seen in the present paper.

A third approach for fuzzy modeling is based on the theory of fuzzy relational equations [7, 17]. This kind of model, called the fuzzy relational model, can be viewed as a simplification of the linguistic models, since a set of fuzzy rules can be written as a fuzzy relation in a relational equation [27]. The advantage of this simplification is that in linguistic models the fuzzy relation is derived from the aggregation of a rule set whose linguistic terms must be determined, whereas in relational models, the fuzzy relation is only a matrix (called the fuzzy relational matrix) to be estimated. However, unlike the rule set of linguistic models, a fuzzy relation in its original form has no clear linguistic meaning. This problem was addressed by Tong [24] and others [5, 6] in their studies of the properties of decomposability of fuzzy relations used to derive fuzzy rules from given (previously estimated) relations. Unfortunately, the conditions necessary for the derivation of the rules are much too restrictive for practical application, especially when the fuzzy relations are estimated using noisy data. Moreover, even when such conditions are satisfied, there is no systematic way of deriving a linguistically clear and consistent rule set.

A partial solution for the problem of linguistic interpretation in relational
models was provided by Pedrycz [16], when he proposed the utilization of the fuzzy discretization technique [26] for data representation in relational equations. This technique has been widespread in the field of fuzzy relational models, especially because it can provide a significant reduction in their dimensions. When using fuzzy discretization, each element of a relational matrix comes to be seen as a measure of possibility of a linguistic rule relating specific input and output reference fuzzy sets [16, 8]. In other words, if $A_i$ and $B_j$ are reference fuzzy sets of the variables of input ($x$) and output ($y$), respectively, then the element $R_{i,j}$ of a relational matrix $R$ relating these variables is interpreted as a rule of the format “IF $x$ is $A_i$, THEN $y$ is $B_j$ with possibility $R_{i,j}$”. From this point of view, fuzzy relational models seem to be a special kind of linguistic model which takes into consideration all possible rules relating input and output fuzzy sets, each rule weighted by the value of the respective relational matrix element. However, it is shown in the present paper that this interpretation is not adequate when specific fuzzy composition operators are involved. Roughly speaking, it is shown that in models using operators such as the well-known max-t composition, certain elements of the relational matrix can contribute more to the output (i.e., can have more influence on the output) than others of equal or greater value. One can thus affirm that a relational matrix element with a small value may be more representative than another with a large value and, consequently, any individual (element-by-element) analysis may give inadequate results.

The use of the concept of representability of a relational matrix element provides a new paradigm for the linguistic interpretation of fuzzy relational models. This concept expresses how much an element of a relational matrix can contribute to the model output and will be investigated in this paper as a tool for the extraction of linguistic information from fuzzy relational models. Specifically, this linguistic information is derived in the form of a set of consistent (non-conflicting) fuzzy rules, i.e., a set with no two rules having the same antecedent yet distinct consequents.

Effective quantitative and qualitative modeling can be achieved by combining the proposed methodology for the extraction of linguistic knowledge with efficient algorithms for fuzzy relational model identification. An algorithm to provide either analytical or numerical solutions, depending on user requirements, is proposed here. When linguistic knowledge is not required, an analytical solution can be derived for a given data set (off-line identification). Also, the Recursive Least Squares (RLS) method [10] can be used for on-line identification. When linguistic knowledge is demanded, however, interpretability constraints are introduced into the optimization problem and quadratic programming can be used to solve it.

The paper is organized as follows. In Section 2, fuzzy relational models are reviewed briefly. The methodology for the extraction of linguistic knowledge from such models is proposed in Section 3 and the identification algorithm is derived in Section 4. The strategies proposed are evaluated in Section 5 using numerical examples and, finally, the conclusions are presented in Section 6.
2 Fuzzy Relational Models

Fuzzy relational models follow the standard fuzzy system structure composed of input and output interfaces and a processing block. The processing block is constituted by a fuzzy relational equation [17], as follows:

\[ Y = X_1 \cdot X_2 \cdot \ldots \cdot X_n \cdot R \quad (1) \]

where \( Y = [Y_1 \ldots Y_c]^T \) and \( X_i = [X_{i1} \ldots X_{ci}]^T (i = 1, \ldots, n) \) are fuzzy representations of the numerical (nonfuzzy) output \( y \) and inputs \( x_i \), respectively, \( R (c_1 \times \cdots \times c_n \times c_0) \) is the fuzzy relational matrix and “\( \cdot \)” denotes the fuzzy composition operator.

Equation (1) can be simplified as

\[ Y = \Omega \cdot R \quad (2) \]

where \( \Omega (c_1 \times \cdots \times c_n) \) is the Cartesian product of the fuzzy inputs,

\[ \Omega = X_1 \times X_2 \times \cdots \times X_n \quad (3) \]

which is generically defined as the product of a triangular norm (t-norm) over the cross product space of the fuzzy inputs.

Both static and dynamic systems can be represented as in Equation (2). Given a set of input-output data pairs, the identification problem is then to find a matrix \( R \) such that this equation is (completely or approximately) satisfied for these data. According to the fuzzy discretization concept [26, 16], the input and output data can be represented (fuzzified) using reference fuzzy sets, as follows:

\[ X_i = [X_{i1}(x_i) \leq \ldots \leq X_{ci}(x_i)]^T \quad (4) \]

\[ Y = [Y_1(y) \leq \ldots \leq Y_{c_0}(y)]^T \quad (5) \]

where \( X_{ih} \) is the \( h \)-th reference fuzzy set in the \( i \)-th input interface and \( Y_j \) is the \( j \)-th reference fuzzy set in the output interface. The reference fuzzy sets \( X_{ih} \) and \( Y_j \) are defined over the universes of discourse \( X_i \) and \( Y \) of \( x_i \) and \( y \), respectively.

The fuzzy composition operator makes an important role in Equation (2). This equation can be rewritten pointwisely for a specific operator. In this case, although \( \Omega \) is an \( n \)-dimensional matrix, it can also be represented as a vector \((l \times 1)\). In the same way, \( R \) can also be expressed as a 2-dimensional matrix, instead of an \((n+1)\)-dimensional one. For example, using the max-t composition (which is the most commonly used composition operator in fuzzy systems), the relational model in (2) is rewritten as:

\[ Y_j = \bigvee_{m=1}^t \Omega_m \tau R_{m,j}, \quad j = 1, \ldots, c_0 \quad (6) \]
where “∨” stands for the max operator, \( \tau \) is a t-norm and \( l = c_1 \cdot c_2 \cdots c_n \).

Instead, if the averaging-product composition [14, 15] is used, then the following equation holds

\[
Y_j = \frac{1}{l} \sum_{m=1}^{l} \Omega_m R_{m,j}, \quad j = 1, \cdots, c_0
\]  

(7)

and \( \Omega \) in (3) becomes the Kronecker product of the fuzzy inputs,

\[
\Omega = X_1 \otimes X_2 \otimes \cdots \otimes X_n
\]  

(8)

Both max-t and averaging-product compositions will be considered in the subsequent discussion.

3 Knowledge Extraction

Consider the generic relational model in (1) as well as the vector \( x = (x_1, \cdots, x_n) \) containing the nonfuzzy inputs of the model, such that \( x \in X \) where \( X = X_1 \times \cdots \times X_n \) is the \( n \)-dimensional space of the input variables.

**Definition 1:** The portion of the output \( Y \) (for a specific input vector \( x \)) related to each element \( R_{m,j} \) of the relational matrix is defined as the contribution \( C \) of \( R_{m,j} \), denoted by \( C(R_{m,j}, x) \).

The contribution of a relational matrix element indicates how much this element contributes to the model output for a given input vector \( x \). So, this concept applies to specific values of \( x \). A measure of contribution over the complete input space is introduced in the following definition:

**Definition 2:** The representability \( \mathcal{R} \) of an element \( R_{m,j} \) is defined as the sum of the contributions \( C \) of this element for all possible values of the input variables \( x_1 \in X_1, \cdots, x_n \in X_n \), i.e. for all \( x \in X \), as follows:

\[
\mathcal{R}(R_{m,j}) = \int_{x \in X} C(R_{m,j}, x) \, dx
\]  

(9)

As outlined in Section 1, the representability of a relational matrix element indicates how much this element can contribute to the model output.

**Definition 3:** The matrix \( \Theta \) containing the representability measures of a relational matrix \( R \) is defined as the representability matrix of \( R \).

**Remark 1:** According to Definition 1 and Equation (6), the contribution \( C \) of each element \( R_{m,j} \) of a model with max-t composition (for a specific \( x \)) is given by
Consider a single-input/single-output fuzzy relational model with max-
product composition (the t-norm $\tau$ is the algebraic product) and normalized
variables of input $x \in X$ and output $y \in Y$ such that $X = Y = [0, 1]$. Both the
input and output variables have two reference fuzzy sets each (labeled “Small”
and “Big”), which are defined over their universes of discourse $X$ and $Y$, as
shown in Figure 1.

Let the relational matrix of the model be as follows:

$$R = \begin{bmatrix} 0.5 & 0.5 \\ 0.1 & 0.9 \end{bmatrix}$$

(11)

Since each element $R_{m,j}$ relates the reference fuzzy sets $X_m$ and $Y_j$, then us-
ing the conventional approach for the interpretation of fuzzy relational models
discussed in Section 1 leads to the derivation of the following fuzzy rules:

- **IF** $x$ is *Small* **THEN** $y$ is *Small* (possibility 0.5)
- **IF** $x$ is *Small* **THEN** $y$ is *Big* (possibility 0.5)
- **IF** $x$ is *Big* **THEN** $y$ is *Small* (possibility 0.1)
- **IF** $x$ is *Big* **THEN** $y$ is *Big* (possibility 0.9)

Based on the analysis which generated the rule set presented above, one
could interpret the last two rules simply as “**IF** $x$ is *Big** THEN $y$ is *Big**” since
the possibility measure of this rule is greater than that of its competing rule
[16]. But what about the first two rules? These two rules cannot be clearly in-
terpreted since they are (linguistically) in conflict\(^1\). The problem arises because
the conventional linguistic interpretation of relational models discussed up to
here analyzes the relational matrix elements individually instead of considering
the matrix as a whole.

Now, consider the measures of contribution and representability given by
(10) and (9), respectively. Since the model has a single input with two reference
fuzzy sets, then $\Omega = [X_1(x) X_2(x)]^T$ ($l = 2$). Therefore, from the conditions
stated in the beginning of the example (i.e. $X = [0, 1]$ and $\tau =$ algebraic
product), Equations (10) and (9) are rewritten as

$$C(R_{m,j}, x) = \begin{cases} \Omega_m \tau R_{m,j}, & \text{if } m = \arg \max_{i=1,\ldots,l} \Omega_i \tau R_{i,j} \\ 0, & \text{otherwise} \end{cases}$$

(12)

$$R(R_{m,j}) = \int_0^1 C(R_{m,j}, x) \, dx$$

(13)

Furthermore, since the membership functions of the reference fuzzy sets in
Figure 1 are $X_1(x) = 1 - x$ and $X_2(x) = x$, it is easy to verify using (11) that
$\mathcal{X}_1(x)R_{1,1} \geq \mathcal{X}_2(x)R_{2,1}$ if and only if $x \leq \frac{5}{6}$ and $\mathcal{X}_1(x)R_{1,2} \geq \mathcal{X}_2(x)R_{2,2}$ if and only if $x \leq \frac{1}{2\pi}$. Hence, Equation (12) yields

$$
C(R_{1,1}, x) = \begin{cases} 
\mathcal{X}_1(x)R_{1,1}, & \text{if } x \leq \frac{5}{6} \\
0, & \text{otherwise}
\end{cases}
$$

$$
C(R_{2,1}, x) = \begin{cases} 
\mathcal{X}_2(x)R_{2,1}, & \text{if } x \geq \frac{5}{6} \\
0, & \text{otherwise}
\end{cases}
$$

$$
C(R_{1,2}, x) = \begin{cases} 
\mathcal{X}_1(x)R_{1,2}, & \text{if } x \leq \frac{1}{2\pi} \\
0, & \text{otherwise}
\end{cases}
$$

$$
C(R_{2,2}, x) = \begin{cases} 
\mathcal{X}_2(x)R_{2,2}, & \text{if } x \geq \frac{1}{2\pi} \\
0, & \text{otherwise}
\end{cases}
$$

Using the equations above as well as Equations (11) and (13), the representability of each element of the relational matrix can be computed as illustrated below for $R_{1,1}$:

$$
\mathcal{R}(R_{1,1}) = \int_{0}^{\frac{5}{6}} (1 - x) 0.5 \, dx = 0.2431
$$

The representability measures of the other elements are computed in the same way (mutatis mutandis), yielding the following representability matrix $\Theta$:

$$
\Theta = \begin{bmatrix}
\mathcal{R}(R_{1,1}) & \mathcal{R}(R_{1,2}) \\
\mathcal{R}(R_{2,1}) & \mathcal{R}(R_{2,2})
\end{bmatrix} = \begin{bmatrix}
0.2431 & 0.1467 \\
0.0153 & 0.3926
\end{bmatrix}
$$

This matrix shows that although the elements $R_{1,1}$ and $R_{1,2}$ in (11) have the same value, $R_{1,1}$ is significantly more representative than $R_{1,2}$.

It is proposed here that linguistic knowledge can be extracted by selecting the rule related to the most representative element of each row of the relational matrix. In the example above, the following rules would be extracted:

\begin{itemize}
\item \text{IF } x \text{ is Small THEN } y \text{ is Small}
\item \text{IF } x \text{ is Big THEN } y \text{ is Big}
\end{itemize}

It is important to note that the derivation of analytical solutions for the measures of representability given by Equation (9) (which involves multiple integrals for multi-variable models) may be difficult. The level of difficulty depends on the composition operator, the number of reference fuzzy sets and the shape of their membership functions, as well as the dimension of the model input space. However, the use of an adequate data set (homogeneously distributed in the input space) makes a reasonable approximation of (9) possible through the approximation of the integral by the sum of the contributions computed individually for each input datum. So, given a set of $N$ input data $\{ x(k) \}_{k=1}^{N}$ the representability of an element $R_{m,j}$ can be approximated by

$$
\mathcal{R}(R_{m,j}) = \sum_{k=1}^{N} C(R_{m,j}, x(k))
$$

7
An interesting alternative is a judicious choice of the composition operator in Equation (2). For instance, if the averaging-product composition is used, as in Equation (7), then a simple but very useful feature related to the extraction of linguistic knowledge from relational models is verified, as follows.

**Remark 2:** Based on Definition 1 and Equation (7), the contribution of each element \( R_{m,j} \) of a model with averaging-product composition (for a specific input vector \( x \)) is

\[
C(R_{m,j}, x) = \frac{1}{l} (\Omega_m R_{m,j})
\]

In this specific case, the contribution and (consequently) the representability of an element \( R_{m,j} \) are linearly (directly) proportional to the value of the element. Thus, elements of the same row of the relational matrix having equal values provide equal contributions and have equal representabilities. Therefore, the rule set can be extracted simply by selecting the rule related to the largest element of each row of the relational matrix. The computation of the representability matrix \( \Theta \) is not necessary. This means that, when the relational model is formulated using the averaging-product composition, the method proposed to extract fuzzy rules based on the representability matrix coincides with the conventional approach introduced by Pedrycz [16] for models with max-t composition, discussed in Example A.

**Example B**

Consider the same conditions given in Example A, but a relational matrix given by

\[
R = \begin{bmatrix}
0.3 & 0.7 \\
0.8 & 0.2
\end{bmatrix}
\]

and the averaging-product composition. As shown in Remark 2, the representability measures of the relational matrix elements are directly proportional to their values (whenever the averaging-product composition is used). So, the elements \( R_{1,2} \) and \( R_{2,1} \) in (14) are the most representative elements of the rows 1 and 2 of the matrix, respectively. Bearing in mind that each element \( R_{m,j} \) relates the reference fuzzy sets \( X_m \) and \( Y_j \) (see Figure(1)), the following fuzzy rules can be extracted without any computation of representability values:

- IF \( x \) is Small THEN \( y \) is Big
- IF \( x \) is Big THEN \( y \) is Small

A natural question which may arise is: what to do when the most representative elements of a specific row of the relational matrix have similar values of representability? (This would be the case of the model in Example A with the averaging-product composition, instead of the max-product one). The answer is that such a situation is unlikely to occur for accurate models, since a system can only be modeled accurately if a consistent set of fuzzy rules is used. If conflicting rules are introduced, then the model tends to be inaccurate, involving
extensive interpolation and low specificity. To investigate this idea a method for the identification of fuzzy relational models is developed in the next section. The models are based on the averaging-product composition operator since it simplifies the knowledge extraction procedure, as shown in Example B. The solution of the identification problem is also simplified if this operator is chosen.

4 Relational Model Identification

The optimization criterion used here for the identification of fuzzy models is based on one of the findings presented in [3, 4]. This result is reported below.

Consider the estimated (nonfuzzy) output of the relational model $\tilde{y}$, the actual (measured) output of a real system $y$ and their fuzzy representations $\tilde{Y}$ and $Y$, respectively, such that

$$ Y = \mathcal{L}(y) \quad (15) $$

$$ \tilde{y} = \mathcal{N}(\tilde{Y}) \quad (16) $$

where $\mathcal{L}$ and $\mathcal{N}$ represent the fuzzification and defuzzification mappings, respectively. Also, let the input and output interfaces of the model be constructed as optimal interfaces [13] (so-called lossless encoding/decoding mechanisms [19]), meaning that any numerical value in their universes of discourse can be completely recovered after a fuzzification-defuzzification sequence through the interface, i.e.

$$ \forall \ a \in A : \mathcal{N}(\mathcal{L}(a)) = a \quad (17) $$

Proposition 1

If the mappings $\mathcal{L}$ and $\mathcal{N}$ in (15) and (16) are implemented through an optimal interface, then the equality $\tilde{Y} = Y$ between the fuzzy outputs of the model and the system results in the equality $\tilde{y} = y$ between their respective nonfuzzy outputs.

Proof: If $\tilde{Y} = Y$, then Equation (16) can be rewritten as

$$ \tilde{y} = \mathcal{N}(\tilde{Y}) \quad (18) $$

Substituting (15) into (18) results in

$$ \tilde{y} = \mathcal{N}(\mathcal{L}(y)) \quad (19) $$

Then, the equality $\tilde{y} = y$ is derived from Equation (19) and the optimal interface concept (17).

The result above means that an approach for identification of fuzzy relational models can be derived from the minimization (over the relational matrix $R$) of
a certain measure of distance between the fuzzy outputs of a given system and
the corresponding model, as follows:

$$\min_R J(Y, \tilde{Y})$$

where $J$ is a generic distance measure (objective function). For a set of $N$
in-put/output data pairs, i.e. $\{(x(k), y(k))\}_{k=1}^N$, the following objective function

$$J = \frac{1}{2} \sum_{j=1}^{c_0} \sum_{k=1}^N (Y_j(k) - \tilde{Y}_j(k))^2$$

(20)

Using the averaging-product composition, the fuzzy output of the model $\tilde{Y}$
is given pointwisely by

$$\tilde{Y}_j(k) = \frac{1}{l} \sum_{m=1}^l \Omega_m(k) R_{m,j}, \quad j = 1, \cdots, c_0$$

(21)

From Equation (21) and after some algebraic manipulation of (20), the op-
timization problem becomes

$$\min_{\phi} J = \frac{1}{2} \phi^T H \phi + f^T \phi + g$$

(22)

where $\phi \in \mathbb{R}^{lc_0 \times 1}$, $H \in \mathbb{R}^{lc_0 \times lc_0}$, $f \in \mathbb{R}^{lc_0 \times 1}$ and $g \in \mathbb{R}$ are given by

$$\phi = \text{vec}(R) \triangleq [R_{1,1} \cdots R_{l,1} \cdots R_{1,c_0} \cdots R_{l,c_0}]^T$$

$$H = \frac{1}{l^2} \begin{bmatrix}
\sum_{k=1}^N \Omega(k)\Omega^T(k) & 0 & \cdots & 0 \\
0 & \ddots & \vdots & \\
\vdots & \ddots & \ddots & 0 \\
0 & \cdots & 0 & \sum_{k=1}^N \Omega(k)\Omega^T(k)
\end{bmatrix}$$

$$f = -\frac{1}{l} \begin{bmatrix}
\sum_{k=1}^N Y_1(k)\Omega^T(k) & \cdots & \sum_{k=1}^N Y_{c_0}(k)\Omega^T(k)
\end{bmatrix}^T$$

$$g = \frac{1}{2} \sum_{k=1}^N Y^T(k) Y(k)$$

10
Problem (22) is solved by the calculation of
\[ \frac{\partial J}{\partial \phi} = H \phi + f = 0 \]

If the matrix $H$ is non-singular\(^3\), then the solution is
\[ \phi = -H^{-1} f \]  
(23)

This solution can also be computed recursively based on the classic recursive least squares formulation \([10]\), without any matrix inversion being explicitly necessary. The steps are summarized below:

\[ \phi_j(k+1) = \phi_j(k) + \frac{P(k)\Omega(k+1)\left(l Y_j(k+1) - \Omega^T(k+1)\phi_j(k)\right)}{1 + \Omega^T(k+1)P(k)\Omega(k+1)} \]  
(24)

\[ P(k+1) = P(k) - \frac{P(k)\Omega(k+1)\Omega^T(k+1)P(k)}{1 + \Omega^T(k+1)P(k)\Omega(k+1)} \]  
(25)

where $\phi_j = [R_{1,j} \cdots R_{l,j}]^T$ (for $j = 1, \cdots, c_0$) and $k = 0, \cdots, N - 1$. Usually, the initial values are set as $\phi_j(0) = 0$ and $P(0) = \alpha I$, where $I$ is the identity matrix ($l \times l$) and $\alpha$ is a scalar.

Theoretically, the elements of the relational matrix $R$ should belong to the interval $[0, 1]$. This is a necessary condition for the linguistic interpretation based on the conventional approach, that is, when the elements of $R$ are viewed as possibility measures. To guarantee this, the constraints $0 \leq \phi \leq 1$ must be introduced into Problem (22), which in turn can be solved using quadratic programming \([1]\). Unfortunately, these constraints may lead to poor solutions. When the linguistic interpretation based on the representability matrix is used, however, only the non-negativity constraints $\phi \geq 0$ are mandatory in (22), because the concept of contribution (Definition 1) makes no sense for negative values only. In this case the optimization problem becomes

\[ \min_{\phi} J' = \frac{1}{2} \phi^T H \phi + f^T \phi \]  
(26)

where $J'$ is the portion of $J$ which depends on $\phi$. The formulation above can lead to efficient solutions for the identification problem and allows the linguistic interpretation of the resulting models. This is illustrated by numerical examples in Section 5. Before proceeding, the modeling procedure is summarized below.

4.1 Modeling Procedure

1. Select a set of $N$ input/output data pairs from the system: \[ \{ x_1(k), \cdots, x_n(k), y(k) \}_{k=1}^{N}. \]
2. Define the reference fuzzy sets for the input and output variables and their respective linguistic labels.

3. Derive the set of fuzzy data \( \{ \Omega(k), Y(k) \}_{k=1}^{N} \) from Equations (4), (5) and (8).

4. Identify the relational matrix of model (21): If linguistic knowledge is demanded, then compute the matrix by solving Problem (26) and extract the rule set by taking the fuzzy rules related to the largest element of each row of the matrix, as shown in Example B of Section 3. Otherwise, compute the matrix by using either Equation (23) or Equations (24) and (25).

5 **Numerical Examples**

5.1 **Example 1**

Consider the discrete-time dynamic fuzzy system \( y(k) = \mathcal{F}(y(k-1), u(k-1)) \), where \( \mathcal{F} \) is a mapping given by the 25 fuzzy rules shown in Figure 2, which are implemented with the Sum-Product-Gravity reasoning method [11].

In Figure 2, NB, NS, ZE, PS and PB refer to the linguistic terms Negative Big, Negative Small, Zero, Positive Small and Positive Big, respectively. These linguistic terms are labels for the reference fuzzy sets defined over the universes of discourse of each variable, as shown in Figure 3. In this example, the universe of discourse of both input and output variables is the interval \([-1, 1]\) \((u, y \in [-1, 1])\).

The relational model \( Y(k) = X_1(k) \cdot X_2(k) \cdot R \), with nonfuzzy inputs \( x_1(k) = y(k-1) \) and \( x_2(k) = u(k-1) \), is used for modeling this system. Identification is carried out by solving Problem (26). For statistical relevance, 10 separate identification procedures are performed using distinct input-output data sets, each with the input \( u \) being a sequence (500 samples) of steps of random amplitude and a period equal to 10. Each model derived is simulated (one-step-ahead prediction), and the mean squared error (MSE) between the outputs of the system and the model is computed. The arithmetic mean of the resulting MSE’s is \( 9.03 \cdot 10^{-5} \) (best model: \( 6.16 \cdot 10^{-5} \); worst model: \( 1.36 \cdot 10^{-4} \)). The recursive simulation of an intermediate model (MSE = \( 9.89 \cdot 10^{-5} \)) using data other than those involved in its identification is shown in Figure 4, illustrating the efficiency of the proposed identification method.

Once identification is completed, a set of 25 fuzzy rules is extracted from each model. These rule sets are then compared with that of the original fuzzy system. Completely correct rule sets were derived from nine models, and the other yielded only two wrong rules. Consequently, the proposed methodology yielded an accuracy of 99.2%.

The same procedure of identification and extraction of fuzzy rules is then repeated using a new rule set, shown in Figure 5.
The arithmetic mean of the MSE’s generated from the simulation of the 10 resulting models is $8.38 \cdot 10^{-5}$ (best model: $5.75 \cdot 10^{-5}$; worst model: $1.04 \cdot 10^{-4}$). Completely correct rule sets were derived from eight models, with the ninth model leading to the derivation of two wrong rules and the tenth model to one. The results are thus 98.8% correct, again showing the efficiency of the proposed methodology.

5.2 Example 2

In this example, the well-known gas furnace of Box and Jenkins [2] is considered. The system is a furnace in which air and methane are combined to form a mixture of gases containing CO$_2$ (carbon dioxide). The air feed is maintained constant, whereas the methane gas feed rate (cubic feet per minute) can be varied as desired. The data set consists of 296 input/output data pairs, where the output $y$ is carbon dioxide concentration (%CO$_2$) in the mixture of gases and the input $w$ is a linear transformation of the actual input of the system $u$ (methane feed rate), such that

$$u(k) = 0.60 - 0.04 w(k)$$

(27)

The relational model used for the identification of this system uses the structure investigated in [16, 9], i.e. $Y(k) = X_1(k) \bullet X_2(k) \bullet R$, with the inputs being given by $x_1(k) = y(k - 1)$ and $x_2(k) = w(k - 3)$. The reference fuzzy sets and their respective linguistic terms for the input variable $w$ are the same as those used in Example 1 (see Figure 3), but the universe of discourse is the interval $[-3.5, 3.5]$. The reference fuzzy sets of the output variable $y$ are defined in the same manner. However, since the universe of discourse of this variable is the positive interval $[43, 63]$, its linguistic term set is defined as Very Low (VL), Low (L), Medium (M), High (H) and Very High (VH), instead of NB, NS, ZE, PS and PB, respectively.

The system is identified by solving Problem (26). After identification, the relational model is simulated resulting in MSE = 0.19 in one-step-ahead prediction. Then, a set of fuzzy rules is extracted from the model. The rules are given in Figure 6.

Although the available data do not represent the system dynamics completely and the model does not rely on any (relevant) second order information, the extracted rules illustrate the basic behavior of the system. One way of observing this basic behavior is by means of linguistic trajectory analysis, as shown below. In the present example, slowly increasing and decreasing linguistic sequences of the input variable are used, based on the characteristics of the sequence of input data used for the identification of the system. For example, considering a decreasing linguistic sequence for the input and the initial condition “$w(k-3)$ is PS and $y(k-1)$ is VL,” the linguistic trajectory based on the rule set shown in Figure 6 would be the following:
Using this kind of analysis makes it possible to observe the fundamental characteristic of the system (see Figure 7). The trajectories indicated in Figure 7 show that when the input increases (downward arrows) the output decreases. On the other hand, when the input decreases (upward arrows) the opposite occurs. Since the input \( w \) varies inversely with the actual input \( u \) of the system obtained from Equation (27), the predominant relation between \( u \) and \( y \) is considered to be positive (and delayed). Roughly speaking, this means that when the methane feed rate increases, the reaction in the gas furnace also does so (with more methane being burned) and, as a consequence, after the period of time necessary to complete the reaction (the time delay) the carbon dioxide concentration also increases.

Finally, to validate the rule set extracted, a linguistic (rule-based) model is implemented from these rules in combination with the Sum-Product-Gravity reasoning method. The simulation of this linguistic model is displayed in Figure 8, where it is shown that the extracted rules represent the dynamics of the system reasonably well. This result illustrates once again the efficiency of the combined approach for identification and extraction of linguistic knowledge from systems using relational models.

5.3 Example 3

The problem of approximation of functions by using fuzzy relational models is addressed here. The function presented in [12] is considered (Figure 9):

\[
y(x) = 3e^{-x^{2}}sin(\pi x) + \eta
\]

where \( \eta \) is a disturbance term. The data used for estimating this function are obtained by taking 300 samples of \( y(x) \) contaminated by Gaussian noise \( \eta \) (mean zero and variance 0.15) with random inputs \( x \) uniformly distributed in the interval \([-3,3]\) (universe of discourse of the input variable).

The reference fuzzy sets and their respective linguistic terms are similar to those defined in Example 1, as shown in Figure 10, but with two new linguistic terms NM and PM, meaning Negative Medium and Positive Medium, respectively. In this case, however, the modal values of the fuzzy sets are set (approximately) at the critical points of the function. This illustrates the use of human knowledge in the development of a fuzzy model, in that the fuzzy sets are chosen to represent specific local behaviors of the system.
As in Example 1, 10 experiments using distinct data sets for model estimation are performed. Estimation is carried out by solving Problem (26). The following rule set was extracted from all the resulting models:

\[
\begin{align*}
\text{IF } x \text{ is NB THEN } y \text{ is ZE} \\
\text{IF } x \text{ is NM THEN } y \text{ is PS} \\
\text{IF } x \text{ is NS THEN } y \text{ is NB} \\
\text{IF } x \text{ is PS THEN } y \text{ is PB} \\
\text{IF } x \text{ is PM THEN } y \text{ is NS} \\
\text{IF } x \text{ is PB THEN } y \text{ is ZE}
\end{align*}
\]

By looking at Figure 9, it is easy to see that the extracted rules are quite realistic. However, although the models are linguistically consistent, they are not quantitatively accurate (the arithmetic mean of the MSE’s generated from the simulation of the 10 models is 0.222; best: 0.219; worst: 0.226). This is a consequence of various facts, the most important being that the number of fuzzy sets is not large enough to achieve accurate performance. Other causes of the inaccuracy include the effect of the interpretability constraints $\phi \geq 0$ in (26), which restrict the solution of the estimation problem, and the fact that the input and output interfaces implemented through the reference fuzzy sets shown in Figure 10 are not optimal. However, since the linguistic knowledge has already been extracted, another model can be estimated without regard for linguistic meaning. In this new model, the reference fuzzy sets (8 instead of just 6 for the input variable) are optimized using the $\Sigma$-PAFIO algorithm presented in [13] (see Figure 11). The reference fuzzy sets of the input variable were optimized within the subinterval [-2,2] of the universe of discourse, which led to their concentration in this region of the domain (where the function has complex behavior), thus reducing the number of fuzzy sets needed to approximate the function with the desired degree of accuracy.

Once again, 10 experiments using distinct data sets for model estimation are performed. The models are derived using the analytical (unconstrained) solution provided by Equation (23). Using other data for model validation, obtained from 300 noise free samples of $y(x)$ with the input $x$ equally spaced in the interval [-3,3], the arithmetic mean of the MSE’s between $y(x)$ and the model output for the 10 models identified is 0.0049 (best model: 0.0038; worst model: 0.0066). The performance of an intermediate model (MSE=0.0051) portraying its accuracy is illustrated in Figure 12. This result can clearly be improved, however, by increasing the number of reference fuzzy sets. For example, repeating the experiments using 10 fuzzy sets for the input instead of 8, the arithmetic mean of the MSE’s between $y(x)$ and the model output for the 10 models identified is reduced to 0.0017 (best model: 0.0008; worst model: 0.0028).

Although it is often possible to obtain a unique model being at the same time accurate and linguistically interpretable (as shown in Examples 1 and 2), the present example has shown that it is not always possible. When it is not, however, it may still be possible to obtain accuracy and linguistic knowledge using different models with the same topology and similar methodologies.
6 Conclusions

The present paper has addressed the problem of linguistic interpretation of fuzzy relational models by proposing a methodology for the extraction of linguistic knowledge from systems using such models. In this methodology, the linguistic knowledge is extracted in the form of non-conflicting fuzzy rules. Using relational models, knowledge extraction then follows a natural sequence based on the original ideas of fuzzy sets and fuzzy logic, with the a priori definition of linguistic terms and the a posteriori estimation of the predominant relations existing between these terms. A new algorithm for the identification of relational models has also been proposed. This algorithm can provide analytical or numerical solutions, depending on user requirements, and includes a recursive version suitable for on-line application. Three numerical examples have been presented to show the effectiveness of the methods proposed here for both quantitative and qualitative modeling.

Acknowledgments

The first author acknowledges FAPESP, the Research Foundation of the State of São Paulo, for fellowship 99/03902-6. The second author acknowledges CNPq, the Brazilian National Research Council, for fellowship 301345184.

References


Footnotes

1. In this specific case, the introduction of an intermediate linguistic term like “medium” could lead to the interpretation “IF $x$ is Small THEN $y$ is Medium”. However, if this new term is considered in the model, the new relational matrix can still present linguistic inconsistencies (in the conventional sense) with rows having similar elements and/or rows forming non-convex fuzzy sets.

2. The mapping $\mathcal{L}$ is given by Equation (5), whereas $\mathcal{N}$ is the weighted average defuzzification.

3. Since $H$ is a block diagonal matrix, it is non-singular if and only if the sub-matrix which forms its diagonal blocks is non-singular.

4. The collection of Gaussian reference fuzzy sets shown in Figure 3 (overlapping at a level of 0.5 and with equally spaced centers) satisfies the conditions of the frame of cognition [18, 19], having clear linguistic meaning.
Figure 1: Reference fuzzy sets of the input and output variables.
Figure 2: Fuzzy rules of the mapping $\mathcal{F}$.

<table>
<thead>
<tr>
<th>$u(k-1)$</th>
<th>NB</th>
<th>NS</th>
<th>ZE</th>
<th>PS</th>
<th>PB</th>
</tr>
</thead>
<tbody>
<tr>
<td>NB</td>
<td>NB</td>
<td>NB</td>
<td>NS</td>
<td>NS</td>
<td>ZE</td>
</tr>
<tr>
<td>NS</td>
<td>NB</td>
<td>NS</td>
<td>NS</td>
<td>ZE</td>
<td>PS</td>
</tr>
<tr>
<td>ZE</td>
<td>NS</td>
<td>NS</td>
<td>ZE</td>
<td>PS</td>
<td>PS</td>
</tr>
<tr>
<td>PS</td>
<td>NS</td>
<td>ZE</td>
<td>PS</td>
<td>PS</td>
<td>PB</td>
</tr>
<tr>
<td>PB</td>
<td>ZE</td>
<td>PS</td>
<td>PS</td>
<td>PB</td>
<td>PB</td>
</tr>
</tbody>
</table>
Figure 3: Reference fuzzy sets.
Figure 4: System output (solid line) and synthetic data of the relational model (dotted line); MSE = 4.37 \cdot 10^{-4}.
Figure 5: Fuzzy rules of the mapping $\mathcal{F}$.

<table>
<thead>
<tr>
<th></th>
<th>NB</th>
<th>NS</th>
<th>ZE</th>
<th>PS</th>
<th>PB</th>
</tr>
</thead>
<tbody>
<tr>
<td>NB</td>
<td>NB</td>
<td>NB</td>
<td>NS</td>
<td>ZE</td>
<td>PS</td>
</tr>
<tr>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>ZE</td>
<td>PS</td>
</tr>
<tr>
<td>ZE</td>
<td>NS</td>
<td>ZE</td>
<td>ZE</td>
<td>ZE</td>
<td>PS</td>
</tr>
<tr>
<td>PS</td>
<td>NS</td>
<td>ZE</td>
<td>PS</td>
<td>PS</td>
<td>PS</td>
</tr>
<tr>
<td>PB</td>
<td>NS</td>
<td>ZE</td>
<td>PS</td>
<td>PB</td>
<td>PB</td>
</tr>
</tbody>
</table>

$u(k-1)$

$y(k-1)$

$y(k)$
Figure 6: Fuzzy rules extracted from the gas furnace data.

<table>
<thead>
<tr>
<th></th>
<th>VL</th>
<th>L</th>
<th>M</th>
<th>H</th>
<th>VH</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>NB</strong></td>
<td>L</td>
<td>L</td>
<td>H</td>
<td>VH</td>
<td>VH</td>
</tr>
<tr>
<td><strong>NS</strong></td>
<td>L</td>
<td>M</td>
<td>M</td>
<td>H</td>
<td>VH</td>
</tr>
<tr>
<td><strong>ZE</strong></td>
<td>L</td>
<td>L</td>
<td>M</td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td><strong>PS</strong></td>
<td>VL</td>
<td>L</td>
<td>M</td>
<td>M</td>
<td>VH</td>
</tr>
<tr>
<td><strong>PB</strong></td>
<td>VL</td>
<td>VL</td>
<td>L</td>
<td>H</td>
<td>H</td>
</tr>
</tbody>
</table>
Figure 7: Fuzzy rules and linguistic trajectories for the gas furnace.
Figure 8: System output (solid line) and prediction generated by the extracted rule set (dotted line).
Figure 9: Function \( y(x) = 3e^{-x^2} \sin(\pi x) \) with no noise.
Figure 10: Reference fuzzy sets for the input (above) and the output (below).
Figure 11: Reference fuzzy sets for the input (above) and the output (below).
Figure 12: Actual noise free function (solid line) and model output (dotted line).