Expert guided integration of induced knowledge into a fuzzy knowledge base

Serge Guillaume, L. Magdalena

Expert guided integration of induced knowledge into a fuzzy knowledge base

Serge Guillaume[1]   Luis Magdalena[2]

[1] Cemagref Montpellier, BP 5095, 34033 Montpellier Cedex, France

[2] ETSI Telecomunicacion, UPM, Ciudad Universitaria, 28040 Madrid, Spain

guillaume@montpellier.cemagref.fr, llayos@mat.upm.es

2005

Abstract   This paper proposes a method for building accurate and interpretable systems by integrating expert and induced knowledge into a single knowledge base. To favor the cooperation between expert knowledge and data, the induction process is run under severe constraints to ensure the fully control of the expert. The procedure is made up of two hierarchical steps. Firstly, a common fuzzy input space is designed according to both the data and expert knowledge. The compatibility of the two types of partitions, expert and induced, is checked according to three criteria: range, granularity and semantic interpretation. Secondly, expert rules and induced rules are generated according to the previous common fuzzy input space. Then, induced and expert rules have to be merged into a new rule base. Thanks to the common universe resulting from the first step, rule comparison can be made at the linguistic level only. The possible conflict situations are managed and the most important rule base features, consistency, redundancy and completeness, are studied. The first step is thoroughly described in this paper, while the second is only introduced.

Keywords   Expert knowledge, fuzzy partitioning, rule induction, rule base fusion, man machine cooperation.

1   Introduction

Fuzzy rule based systems (FRBS) have widely demonstrated their capability for classification, modeling or control. One of the reasons for this success is the ability of fuzzy systems to incorporate human expert knowledge, as well as to express the behavior of the system in an interpretable way for humans. Historically, the first kind of fuzzy rule based systems focused on the ability of fuzzy logic to model natural language (Mamdani & Assilian, 1975). This kind of system offered a high semantic level and a good generalization capability. In a first period of development, most fuzzy systems were expert knowledge-based, where the effort of knowledge extraction was concentrated on defining methods to allow acquiring knowledge from the experts. Unfortunately, the complexity of large systems may lead to an insufficient accuracy in the simulation results.

A second approach, proposed in mid eighties, is defined by the Takagi-Sugeno model (Takagi & Sugeno, 1985). This new type of system was based on automatic learning from data. In fact, this was the beginning of the data-driven design methods. Their main advantage was the numerical accuracy while a major drawback was their *black box* behavior, providing a numerical model, whose coefficients have no meaning for experts.

Finally, in the nineties, the lack of learning capabilities characterizing previous works in fuzzy systems, particularly the Mamdani approach, generated a big interest for the study of fuzzy systems with added learning capabilities. The age of Soft-computing (Zadeh, 1997; Bonissone, 1997) has introduced different kinds of hybrid systems, being the nineties the decade of neuro-fuzzy and genetic-fuzzy systems.

The result of those efforts was the generation of a large number of tools offering an automatic, and relatively easy, design of fuzzy systems. But that produced a side effect, the deviation from the seminal purpose directed towards exploiting the descriptive power of the concept of a linguistic variable (Zadeh, 1973; Zadeh, 1975). Instead, the prevailing research was concentrated on increasing the accuracy as much as possible paying little attention to the interpretability of the final model. During this evolution towards accuracy, the derivation methods were improved, the components to be designed were extended, and new model structures were proposed.

1

The age of expert knowledge-based fuzzy systems was over, and the data driven rule generation methods played the main role of fuzzy system design. Even in the case of merging expert and induced knowledge, the expert knowledge usually had the role of initial knowledge, while the data-driven process was considered as a tuning mechanism where the loose of interpretability was completely disregarded.

However, we should remember the initial philosophy of fuzzy set theory directed to serve the bridge between the human understanding and the machine processing. There is a clear need to recover interpretability. The present paper describes a method for building accurate and interpretable systems by integrating expert and induced knowledge into a single knowledge base, favoring the cooperation between expert knowledge and data.

After this introduction, section 2 describes the two main sources of knowledge, expert and data, and the questions related to interpretability. Then, the overall process of knowledge extraction and integration is analyzed by defining the main components of the structure, and their interactions. Section 4 enters on the question of fuzzy partitioning by considering expert and data induced fuzzy partitions, and their semantic integration. Section 5 will briefly comment the aspects related to fuzzy rule extraction and integration. In section 6, the common fuzzy partition design is illustrated on real world data. Finally some conclusions are presented.

2   Knowledge, data and interpretability

The idea of *Interpretability* is usually related to the ability of a model, or a classifier, to express the behavior of the modeled system in an understandable way. This is a subjective property lying on questions as the model structure, the number of variables and rules, or the dimension and structure of the fuzzy partitions. As shown in (Guillaume, 2001), the use of fuzzy formalism is not sufficient to ensure interpretability of a knowledge base. Three conditions have to be fulfilled. First, semantic integrity should be respected within the partition. Secondly, the number of rules should be small. The third condition is specific to complex systems with a large number of input variables: rules considering only a reduced number of variables per rule, have to be handled. This kind of rules is usually referred to as incomplete rules, but in this case, incompleteness is more related to the idea of compressed information than to the absence of information (that is the case for incompleteness of a rule base).

Some authors use the term of *transparency*. This concept is defined as a property that enables us to understand the influence of each system parameter on the system output. It has been mainly applied to neuro-fuzzy systems (Brown & Harris, 1994) and Takagi-Sugeno models (Setnes, Babuska & Verbruggen, 1998).

A recent work (Riid & Rüstern, 2003) analyzes the existence of two aspects of transparency in fuzzy systems. The first one is related to the readability of rules, based on the use of a moderate number of variables, rules and fuzzy sets, and the avoidance of inconsistencies. There is also a second aspect, low-level transparency, that grows out from conformity between the linguistic layer and the inference function of a fuzzy system, where some of the properties of fuzzy partitions defined in (de Oliveira, 1999) as coverage and distinguishability of membership functions are considered.

Our approach will consider both aspects, the question related to interpretability that consider the readability of the fuzzy partitions, and those considering the readability of the rule base, but the present paper will mostly concentrate on the first one. The second question is mostly related to the effect of searching for accuracy and obtaining a rule base with an excessive size, usually caused by a deficient rule base learning process with tendency to generate too many rules. In this kind of situation it is possible to find (Casillas, Cordon, Herrera & Magdalena, 2003) *redundant rules*, which do not contain relevant information and whose actions are covered by other rules; *erroneous rules*, which are wrong defined and distort the FRBS performance; and *conflictive rules*, which perturb the FRBS performance when coexist with others. Besides worsening the accuracy, an excessive number of rules makes difficult to understand the model behavior.

But interpretability is only one of the concerns of the present paper, the main interest is concentrated on the integration of expert-knowledge and induced knowledge within that framework of no interpretability lose. And this interest for integrating expert and induced knowledge is based on the fact that both kinds of knowledge are different not only in the method of extraction, but also in their intrinsic characteristics.

We detail the proper characteristics of these two kinds of knowledge in order to show how complementary they are and to underline that their cooperation should be benefit for intelligent system performance.

Data, the extensive form of induced knowledge, are basically pieces of positive evidence. The most important data

characteristic is incompleteness: A data sample cannot pretend cover all the possible situations, especially in the case of complex systems. This is why, it is a cautious attitude to consider any sample as partial.

Nevertheless, the data should be processed using reasoning mechanisms to state a confidence level in the individual items and/or to produce new, more general, information.

Induction is the process of estimating the validity of observations of part of a class of facts for a proposition about the whole class. Rule induction consists in generating general rules from a sample.

Induced rule quality highly depends on the training set characteristics: the more representative of the whole class the examples are, the more general the induced rules behave. The representativeness can be assessed by checking the variable distributions, are the sample distributions similar to the real ones ?, and their combination, do they cover the whole input space ?

Induced rule quality also depends on the induction methods itself: to cooperate with expert knowledge, rule induction should express the hidden knowledge in data in an interpretable way (Guillaume, 2001). The goal is not only building an accurate system, in the machine learning sense, but to yield new pieces of knowledge the expert may validate or not.

The other kind of knowledge we are interested in is expert knowledge.

According to the dictionary, an expert is "a person with great knowledge, skill, and experience in a particular field" (Larichev, 2002). (Cornelissen, van der Bergand W.J. Koops & Kaymak, 2003) underline that an expert is a person whose knowledge in a specific domain is obtained gradually through a period of learning and experience, what can be called practice. For (Roy, 2000) practice is nothing but a process of generating and collecting supporting examples in order for the brain to learn.

Expert knowledge acquisition is identified as a difficult task.

Within the multiple model approach proposed in (Shreiber, Wielinga & Breuker, 1993) knowledge acquisition is an iterative process. In this framework, the expert system is not reduced to an inference engine: it is viewed as a model which exhibits a behavior. This behavior, which may coincide or not with the expert one, result of the cooperation of various models, one of them being the model of expertise. The expert system helps in making expert knowledge more explicit and coherent by performing tasks an expert cannot: store a great number of hypothesis in a short term memory, explicit its reasoning process, . . .

But, the most usual, and cheapest, way for collecting expert knowledge is asking the expert to write rules. This transfer is obviously limited. There is evidence that an essential part of expert knowledge in different professional fields is located on the unconscious level and could not be discovered by questioning the expert (Larichev, 2002).

However, experts know the main trends of the most influential variables in the system and are able to describe their behavior using incomplete rules. Expert rules are universal rules due to the cumulated experience they are based on and present a high level of generalization. Dealing with complex systems, they face two difficulties: useful variable identification and variable interaction formalization.

As recalled above, data exhibits complementary characteristics. Induced rules aim to be general rules but are not universal rules. This intrinsic limitation relies on the finished size of the training set. Their generalization ability highly depends on the training set quality, examples with default or failure are usually not available. On the other hand, data may contain a good image of variable interaction.

To take advantage of both kinds of knowledge, one should use them in their respective domains. The cooperation will stand at the two main steps of the fuzzy rule based system design. First, expert knowledge and data cooperate for defining a common partitioning for each of the input and output variables. Secondly, expert rules and induced rules are generated according to the previous common fuzzy universe. Then, induced and expert rules have to be merged into a new rule base. Thanks to the common universe resulting from the first step, rule comparison can be made at the linguistic level only.

3   Process overview

As previously stated, two different kinds of knowledge will be considered: induced knowledge and expert knowledge. This two kinds of knowledge have a different source that is reflected in different methods for extracting it. Neither the induced nor the expert knowledge use to provide a complete and perfect view of the problem to be solved, and that is the reason to focus on the process of integrating both kinds of knowledge. Knowledge integration is not an easy task,

mostly because knowledge representation using a fuzzy approach includes different levels of description, basically the one of linguistic variables and that of fuzzy rules.

In addition to that question, it is important to consider that a wide range of different representations of fuzzy information are considered in FRBS, and in most cases, the knowledge induction methods are tightly connected to a specific representation. Even if we restrict to Mamdani fuzzy controllers, without considering TSK models, it is possible to use descriptive (linguistic) or approximate rules, and we can work on the basis of semantically constrained, or not, fuzzy partitions. At the end, the final result, is a framework too much open to allow an easy integration of knowledge. To cope with this situation, we will superimpose some restrictions that will delineate the extraction and integration process. Most of this restrictions are not only choices to ease the integration, but are mandatory conditions to allow it.

The assumptions are as follows:

1. Descriptive rules based on strong fuzzy partitions will be considered.

2. Linguistic variables are assumed to be the same for induced and expert knowledge, and consequently, the integration of all the available knowledge for partition design will be previous to the rules extraction processes, both expert and induced.

3. The expert is assumed to be able to evaluate induced knowledge, and consequently, the whole integration process will be under expert control.
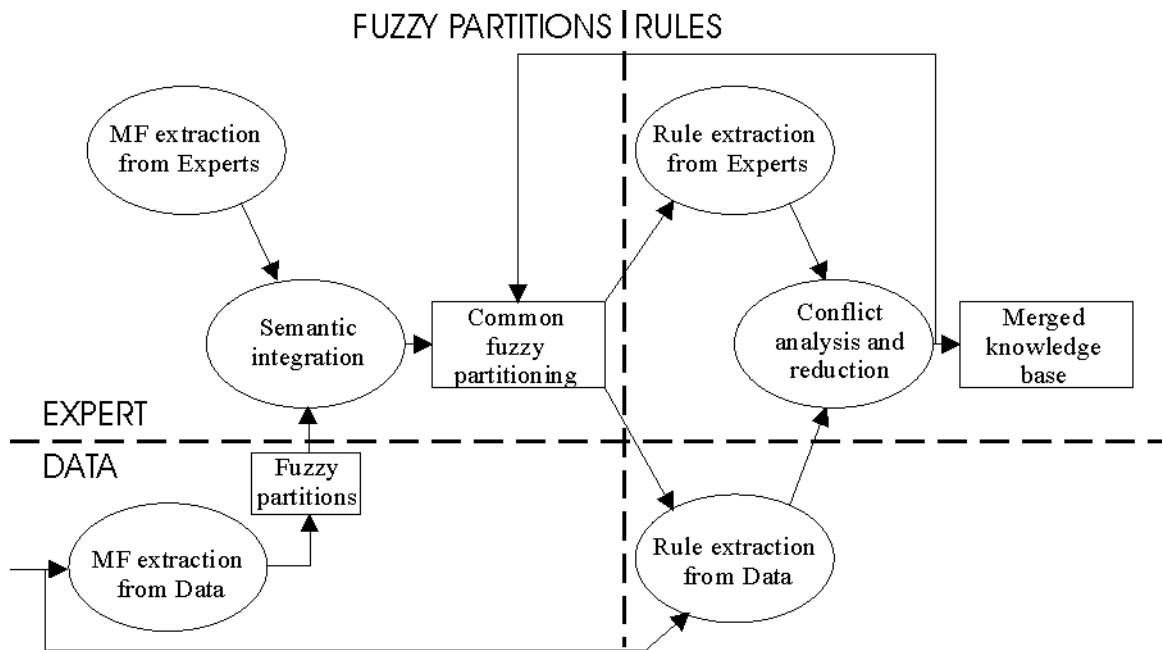

Figure 1: Overall structure of the knowledge extraction process

Having in mind all these restrictions, the extraction and integration process shown in Figure 1 has been defined. As a result of the first two assumptions, the whole process is divided into the left, fuzzy partitions, and right, fuzzy rules, parts. The left part is concentrated on the generation of a common fuzzy partitioning by integrating qualitative expert knowledge and induced partitions with the idea of preserving the semantic integrity. This first part of the process is illustrated in Figure 4 and described in section 4.

Finally, the third assumption produces the division in between the upper, expert, and the lower, data, parts of the figure. The two integration processes, semantic integration and conflicts analysis and reduction, are in the upper part, under the control of the expert.

The feedback, from conflict analysis during rule integration to the definition of the common fuzzy partitioning, only concerns output variables. Some conflicts may lye in an insufficient number of linguistic terms. Output partition refinement will restart the rule generation and merging processes.

The overall structure implies the definition of six main blocks or tasks (represented by ellipses in Figure 1:

- Definition of qualitative partitions by experts

- Extraction of fuzzy partitions from data, with semantic constraints

- Integration of expert and induced fuzzy partitioning

- Extraction of rules by experts

- Extraction of rules from data

- Integration of expert and induced fuzzy rule bases

The four tasks related to extraction are widely analyzed in the literature, and are not central questions to develop the structure. In any case, the choice for a certain approach in any of those tasks must be driven by the general conditions imposed previously, descriptive rules and strong fuzzy partitions. Not all the methods found in the literature fulfill both restrictions, mostly in what refers to knowledge induction methods.

Our objective is to deal with the two integration tasks. The second one consists in analyzing the conflicts which may result from the two rule base merging. It will be introduced in section 5. The first one, whose aim is to design variable partitioning, is detailed in the following section.

4 Defining a common fuzzy partitioning

As previously said, one of the central question of this paper is the integration of expert and induced knowledge into a single knowledge base. This integration will be driven by interpretability criteria. We have also stated that interpretability can be considered at two different levels: fuzzy partitions and fuzzy rules. This section will consider the first level of the process, that of fuzzy partitions. We want to underline the importance of fuzzy partitioning since the possibility of integration will be highly limited if fuzzy partitions do not satisfy certain properties.

The section covers different aspects related to fuzzy partitioning. The first step is to define the idea of interpretability when describing a fuzzy partition (Section 4.1) since it will be quite important in the proposed approach. Then we enter on the design process itself (Section 4.2) by briefly introducing what has been proposed in the literature considering both expert knowledge and data driven approaches. For most of the data driven approaches it is crucial to define evaluation criteria to determine what a good fuzzy partition is. A brief review of those criteria is included in Section 4.3. Finally, once covered all the previous steps of designing and evaluating fuzzy partitions, we are ready to cope with the central aspect of our approach: the semantic integration of expert-based and data-based partitions (Section 4.4).

4.1 The importance of fuzzy partitioning

The readability of fuzzy partitioning is a pre-requisite condition to build an interpretable rule base. The necessary conditions for interpretable fuzzy partitions have been studied by several authors (Ruspini, 1982; de Oliveira, 1999; Glorennec, 1999; Espinosa & Vandewalle, 2000). Let us recall the main points:

- Distinguishability: Semantic integrity requires that the membership functions represent a linguistic concept and different from each other.

- A justifiable number of fuzzy sets.

- Coverage: Each data point, x, should belong significantly, $\mu(x) > \epsilon$, at least to one fuzzy set. ϵ is called the coverage level (Pedrycz, 1993).

- Normalization: All the fuzzy sets should be normal.

- Overlapping: All the fuzzy sets should significantly overlap.

We implement these constraints as follows:

$$\begin{cases} \forall x \sum_{f=1,2,\ldots,m} \mu^f(x) = 1 \\ \forall f \, \exists \, x \, \mu^f(x) = 1 \end{cases} \tag{1}$$

where m is the number of fuzzy sets in the partition and $\mu^f(x)$ is the membership degree of x to the fth fuzzy set. Equation 1 means that any point belongs at most to two fuzzy sets when the fuzzy sets are convex.

Due to their specific properties (Pedrycz, 1994) we choose all fuzzy sets of triangular shape, except at the domain edges, where they are semi trapezoidal. A triangle fuzzy set f is defined by its breakpoints $left^f, c^f, right^f$. Conditions from equation 1 are implemented by choosing fuzzy set breakpoints as shown in figure 2.
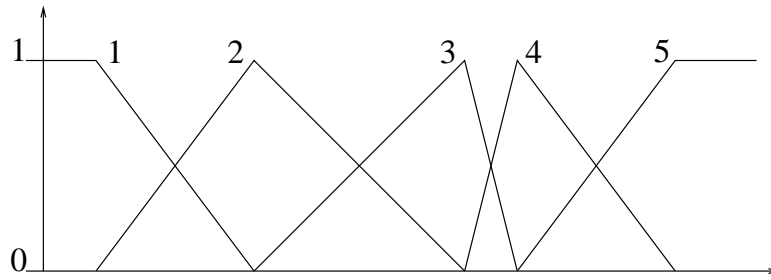

Figure 2: A standardized fuzzy partition

Three contiguous fuzzy sets f, g, h have such boundaries as:

$$\begin{cases} left^g = c^f \\ c^g = left^h = right^f \\ right^g = c^h \end{cases}$$

The objective is to design such highly interpretable kind of partition by integrating expert knowledge and data distribution.

4.2 Fuzzy partition design

This subsection covers the left most parts, initial steps, of Figure 1, i.e., those related to membership functions extraction from both, experts and data. A first part will consider the extraction from experts, while a second one will cover some approaches for membership function design from data.

When considering expert knowledge, we will assume that the expert is able to provide us with, at least, a minimal knowledge about the linguistic variables of the system. For a given variable, experts may define a domain of interest, included in a physical range. Experts make their reasoning using a given, and small, number of linguistic terms. For some of them, they are able to define prototypes: a doctor knows that a typical normal adult temperature is about 37 ^oC. There are several common concepts for membership function definition from expert. The literature establishes different ideas as point estimation, interval estimation and direct rating (Turksen, 1991). Recently, (Cornelissen, van der Bergand W.J. Koops & Kaymak, 2003) a new approach, called the transition interval estimation, has been defined. It is also possible to consider parameterized functions in expert knowledge extraction, where the main task is the adjustment of the parameters by means of statistical methods. Some of them are described in (Medasani, Kim & Krishnapuram, 1998).

In our case, a minimum information from the expert, in membership function definition, will be assumed: only the definition of universes, number of terms, and, sometimes, prototypes of linguistic labels (modal points). This approach is a reduced version of the interval estimation method, reducing the interval to a single point. If the expert is able to generate additional information, it will be also integrated, but we start from that point.

The knowledge base of a FRBS is divided into two main parts, the data base (DB) and the rule base (RB). The data base is usually considered to be made up of the description of the linguistic variables (number, range, granularity, membership functions), and the scaling or normalization function related to those variables. In what concerns the generation from data of the fuzzy partitions, most of the previous works consider the automatic design of the data base as one of the steps in the definition of the overall knowledge base of a fuzzy rule based system. This situation is slightly modified in our approach, since rules will be generated on the basis of fuzzy partitions that integrate both, expert and data-based knowledge. The automatic generation of fuzzy partitions involves either defining the most appropriate shapes for the membership functions, determining the optimum number of linguistic terms in the fuzzy partitions (i.e., the granularity), and/or locating the fuzzy sets into the universe of the variable.

It is possible to define different approaches of the DB design regarding the way of integrating it in the whole knowledge extraction process. Only some of the presented approaches will be applicable in our case.

According to (Casillas, Cordon, Herrera & Magdalena, 2003), the design of the DB may be integrated within the whole design process according to four schemes:

- *Preliminary design* involves extracting the DB a priori by induction from the available data set. This process is usually performed by non supervised clustering techniques (López, Magdalena & Velasco, 1999; Liao, Celmins & Hammell II, 2003), data equalization techniques (Pedrycz, 2001) or by an ascending method based on fuzzy set merging (Guillaume & Charnomordic, 2003). These approaches consider only information related to input variables, and no assumption about the output is made.

- *Embedded design* derives the DB using an embedded basic learning method (Cordón, Herrera, Magdalena & Villar, 2001; Glorennec, 1996; Ishibuchi & Murata, 1996).

- *Simultaneous design* implies that the DB is jointly developed with the derivation of other components such as the rule base in a simultaneous procedure (Homaifar & McCormick, 1995; Krishnakumar & Satyadas, 1995; Magdalena & Monasterio-Huelin, 1997).

- *A posteriori design*, usually called DB *tuning*, involves refining the DB from a previous definition once the remaining components have been obtained (Bonissone, Khedkar & Chen, 1996; Cordón & Herrera, 1997; Karr, 1991; Nauck & Kruse, 1999).

Only the two first approaches are well suited for the framework proposed in this paper since simultaneous learning is not adapted to our *step by step* process. Simultaneous design usually produces a much more complex process where the computational effort is partly useless since only the fuzzy partitions are be considered. A posteriori design could be considered in further steps of the process, but not during the definition of the common fuzzy partitioning. In addition, our interpretability quest requires some specific properties for the partitions. The result is that those clustering techniques generating multidimensional clusters (Liao, Celmins & Hammell II, 2003) cannot be applied, since we need one-dimensional membership functions, obtained by *independently* partitioning the universe of each variable. As a general approach, it is possible to use any one-dimensional optimization technique if it includes some semantic constraints.

In the case of embedded design, when search techniques are used to design the DB, it is possible, to ensure the integrity properties, to include interpretability measures in the objective function, thus guiding the trek to good solutions. Usually, measures of completeness, consistency (Jin, von Seelen & Sendhoff, 1999), compactness, or similarity (Jiménez, Gómez-Skarmeta, Roubos & Babuška, 2001) are considered. At the end of the embedded design process, only the fuzzy partitions are considered as the fuzzy rules will be generated once the common fuzzy partitions have been defined.

These methods, applied to a given data set, lead to very different partitions. The final choice may be guided by numerical criteria.

4.3 Criteria for the evaluation of fuzzy partitions

The evaluation of fuzzy partitions involves two different aspects: linguistic properties and partitioning properties. Linguistic properties are solely related to the shape of the fuzzy sets, and the relations between fuzzy sets defined for the same variable. Their assessment does not involve the data. On the other hand, the partitioning properties refer to the level of matching of the partition with the data from which the partitions have been derived. It is only applicable for partitions derived from data.

The main linguistic properties have been considered at the beginning of this section, being distinguishability, overlapping, normalization, etc. They are guaranteed by the strong fuzzy partition constraints. The present subsection will concentrate on partitioning related properties, i.e., in those related to the way of reflecting in the fuzzy sets, the distribution of data. For this purpose, there are several criteria, some of them related to clustering techniques, that can be applied and will be briefly described.

The first question to be pointed out is that we will only consider those methods relying only on the distribution of data itself, excluding the ones based on an input-output relation. Consequently we focus on those methods that are usually applied in unsupervised clustering, and not in supervised clustering. The earlier methods are those proposed by Bezdek (Bezdek, 1981) as the partition coefficient and the partition entropy. To describe them we will first introduce a common notation. As the criteria are defined to evaluate the matching in between a data distribution and a partition, we will consider the assignment of elements of the data set to each of the elements of the partition (assuming a fuzzy partition). Consequently we will represent the degree of membership of the k-th element of the data set to the i-th element of the fuzzy partition as u_{ik}. The number of elements of the fuzzy partition will be represented by c, while the cardinality of the data set will be n.

With this definition, the partition coefficient is obtained as

$$PC = \frac{\sum\limits_{k=1}^{n}\sum\limits_{i=1}^{c} u_{ik}^2}{n},$$

and the partition entropy as

$$PE = -\frac{1}{n}\left\{\sum_{k=1}^{n}\sum_{i=1}^{c}\left[u_{ik}\log_a(u_{ik})\right]\right\}.$$

Recently, a different index have been proposed by (Chen, 2002)

$$Ch = \frac{1}{n}\sum_{k=1}^{n}\max_{i} u_{ik} - \frac{2}{c(c-1)}\sum_{i=1}^{c-1}\sum_{j=i+1}^{c}\frac{1}{n}\sum_{k=1}^{n}\min(u_{ik}, u_{jk})$$

The three indices described above can be applied to any partition, independently of the derivation method. According to these criteria a good partition should minimize entropy and maximize the coefficient partition and the Chen index.

It is also possible to define some indices that are particularly suited for specific clustering algorithms. A kind of partitioning methods is that of centroid based fuzzy clustering, i.e., techniques where the partition is defined by means of a family of centroids, where each element of the data set is assigned to the different clusters of the partition according to the distance of the data element to the different centroids. A well known validity index for this situation is the Xie-Beni index (Xie & Beni, 1991), to be minimized:

$$XB = \frac{\sum\limits_{k=1}^{n}\sum\limits_{i=1}^{c} u_{ik}^2 \|x_k - v_i\|^2}{n(\min\limits_{i\neq j}\{\|v_i - v_j\|^2\}}.$$

where v_i stands for the $i - th$ cluster center.

The most extended centroid based clustering algorithm is the fuzzy c-means (FCM) one, introduced by (Dunn, 1973), for which some additional indices are defined. A first option is the Xie-Beni index that can be adapted by

replacing u_{ik}^2 with u_{ik}^m, being m the weighting exponent on each membership function. In addition it is possible to use the Fukuyama-Sugeno index (Fukuyama & Sugeno, 1989), to be minimized:

$$FS = \sum_{k=1}^{n} \sum_{i=1}^{c} u_{ik}^m \left(\|x_k - v_i\|_M^2 - \|v_i - \overline{v}\|_M^2 \right),$$

where M is any positive definite symmetric matrix, being the identity and covariance matrices the most commonly used.

For what concerns our approach, any of the partition based index could be applied, on the other side, those integrating the weighting exponent m will not be considered. Consequently, the options are the *partition coefficient*, the *partition entropy*, the *Xie-Beni index* or the *Chen index*. These indices can be used either as optimization criteria to design fuzzy partitions from data or as quality criteria to evaluate induced partitions.

There are other options to encompass the distribution of data with the fuzziness of the partition without using clustering techniques. Some possibilities are to perform a data equalization or to use the *energy* and *entropy* measures of fuzziness (of the fuzzy sets defining the partition) for the given experimental data. The idea of data equalization is applied by (Pedrycz, 2001) when defining

$$P(A) = \int_x A(x)p(x)\mathrm{d}x$$

being A a fuzzy set, and $p(x)$ the probability density function related to the data. Using the previously defined $P(A)$, the equalization condition requires that:

$$P(A_1) = P(A_2) = \ldots = P(A_c) = \frac{1}{c}.$$

If we apply this equation considering the previous notation, the result is:

$$\frac{1}{n} \sum_{k=1}^{n} u_{ik} = \frac{1}{c} \quad \forall i = 1, \ldots, c.$$

A more elaborated but related idea is that of using *energy* and *entropy* measures (Pedrycz & Vukovich, 2002) as part of two indices, to be maximized:

$$\frac{\sum_{k=1}^{n} e(A(x_k))}{\mathrm{card}(\mathrm{supp}(A))}$$

and

$$\frac{\sum_{k=1}^{n}[1 - h(A(x_k))]}{\mathrm{card}(\mathrm{supp}(A))}$$

being e an energy measure of fuzziness, and h an entropy measure of fuzziness. Both, e and h are continuous functions from [0,1] to [0,1], but with different properties. Energy is monotonically strictly increasing with $e(0) = 0$ and $e(1) = 1$, while entropy is monotonically increasing in [0,1/2] and monotonically decreasing in [1/2,1] with $h(0) = h(1) = 0$ and $h(1/2) = 1$.

In (Guillaume & Charnomordic, 2004), the proposed index is based on the homogeneity of the fuzzy set densities. A fuzzy set density, called d_i for fuzzy set i, is defined by the ratio of the fuzzy cardinality, $\sum_{k=1}^{n} u_{ik}$, to the area. The overlapping subareas are excluded, as shown in figure 3, as they are imposed by the standardized fuzzy partition construction. The density homogeneity, σ^{FP}, is defined as the density standard deviation for all the fuzzy sets of the partition:

$$\sigma^{FP} = \sqrt{\frac{1}{c} \sum_{i=1}^{c} (d_i - \overline{d})^2}$$

9

being $\bar{d}$ the mean of the fuzzy set densities. A good, steady, partition is expected to be homogeneous, i.e. to have a small standard deviation.
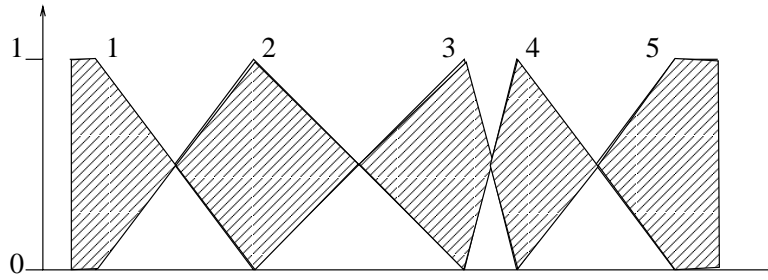

Figure 3: Areas used for computing the fuzzy set densities

4.4 Semantic integration

The agreement between expert-based and data-based partitions must be reached for all the elements defined by the expert: range, granularity and modal points (prototypes).

The information provided by the expert is highly qualitative, being only related to some of the elements previously mentioned. As an example, a typical case will be that of the expert providing only the range and the number of terms that reflects his/her knowledge about the problem under consideration. The task of quantifying that qualitative knowledge will in most cases be left to the data driven parts.

This integration is made variable by variable. It is done under expert control for each of the three steps of the procedure, an overview is given in figure 4. The cases of potential conflict are described to the expert, but at each step the expert makes the final decision.

The underlying assumption is the high level of reliability of expert knowledge: all the pieces of information provided by the expert are considered as almost sure and have priority over knowledge induced from data. On the other hand, as expert knowledge may be questioned by the integration, the expert may update his knowledge all along the procedure.

Data range: If it differs significantly from the range of interest given by the expert, the data set will not help to solve the problem. Otherwise, two situations are possible:

- The expert has defined a part of domain where there is no data available. This may correspond to a part of space where collecting data is expensive or difficult, failures or experiments in extreme condition for instance. In this case, the expert will be asked to specify a modal point in the concerned area.

- Part of data items are out of the expert range. The advice to the expert depends on the data distribution: the items should be considered as outliers or prototypes.

The generic term outliers should be understood in a wide sense, and not restricted to the usual definition, an erroneous data. The data may correspond to a real situation in which the expert is not interested in, or a situation the expert does want to consider to get a simplified model. The starting phase of most industrial continuous processes falls into this category: as it is done only once or twice a year, even it takes several hours or days, it is usually considered as a distinct process.

Outliers detection involves the data distribution, they are expected to be isolated cases, and also a multidimensional analysis: only the whole context may help the expert to make a decision. Once identified, the outliers are removed.

If the out of range items are numerous and concentrated in given areas, the expert has to decide if they are representative of typical cases of interest. Otherwise, they are treated as outliers and removed. As already mentioned,
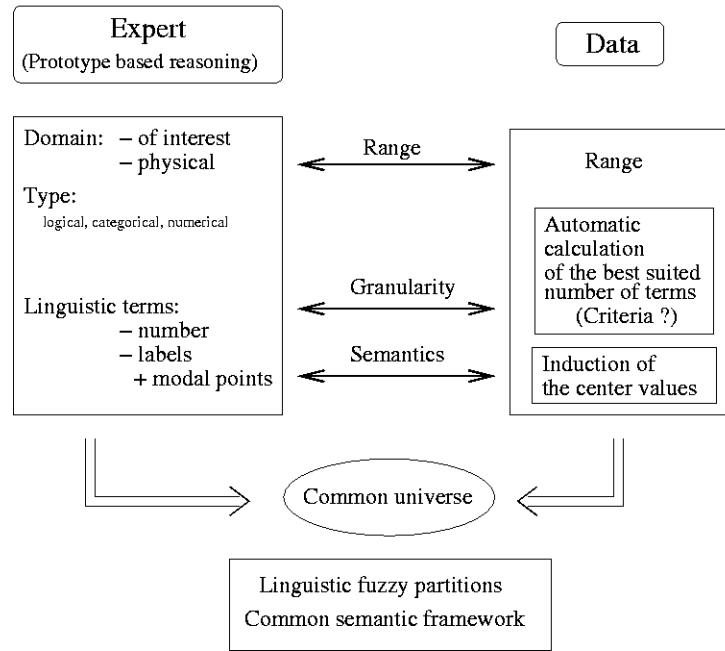
Figure 4: Defining a common universe

the decision can be made by getting a multidimensional view. A positive answer carries two consequences. First, the expert must update the variable range, second it can specify the corresponding modal points.

In the following the cooperation is obviously restricted to the common range, if some modal points have been defined out of the data range, they will serve as additional fuzzy set centers.

Number of terms: According to what has been stated before, the number of terms we are speaking about are all located in the common range. Two options are possible. First, a partition with the number of terms given by the expert can be induced through various ways, see section 4.2. The other option consists in applying an automatic method to determine the best-suited number of terms. The result, which depends on both the method used and the quality criterion, may differ with the number proposed by the expert. The expert has to make the final choice.

Fuzzy set centers: Once an agreement on the number of terms is reached, the final check is only semantical. As the generated partition is a strong fuzzy partition, it is completely defined by the center locations. The checking is limited to their position in the universe, in relation with their associated semantic label: are they possible prototypes according to expert knowledge ? Are the difference with the available modal points acceptable ? As in the previous steps, if needed, the expert makes the final semantical adjustment.

The result of this integration procedure is, for each of the input and output variables, a fuzzy partition designed according to data and carrying the expert knowledge semantic.

5   Rule base generation and integration

As a result of the common fuzzy partition obtained with semantic constrains, the interpretability is ensured in what concerns the linguistic variables. The question now is the derivation of rules without loosing that interpretability.

11

One important question to be considered is that according to Figure 1 the two processes of rules extraction will be complementary at a certain level, i.e., the questions related to interpretability should be analyzed on the whole set of rules once the integration process takes place. As a consequence, some of the *problems* being present in the separate sets of rules, expert and induced, may stay or not in the integrated set. On the other hand, some new conflicts may appear after integration.

The main structural properties to be considered when analyzing a set of rules are completeness, consistency and non redundancy. These questions are *crisp* properties, in the sense that a rule base is or not complete, and two rules are or not consistent or redundant. We can also quantify the level of satisfaction of those properties by defining a level of incompleteness, inconsistency or redundancy. In addition, one important question related to interpretability is the simplicity of the rules base, also called compactness, a property that is mostly qualitative. Simplicity refers to the number of rules in the rule base, and the number of variables that are considered in each rule.

There are different methods for rule extraction that pay attention on some of these properties or qualities. We will not describe those methods that are widely analyzed in the literature. In any case, what is important to consider is what kind of property from a partial rule base will influence or not the integrated rule base, in order to define the key properties of the rule extraction process.

Obviously, we must not take care on completeness until the rule base is integrated, since the main idea of integration is to complete the rule base. On the other hand, inconsistency and redundancy may appear during integration and should be solved on that step. Finally, what is specific of the separate processes of generation, is the compactness or not of the sets of rules. Merging two compact sets of rules, ensuring in the merging process consistency and non redundancy, may enlarge the number of rules but to complete areas where no knowledge is available in one of the rule sets.

As a result of previous comments, the most important question to be considered during rule generation is the compactness of the rules, while consistency and non redundancy should be managed during integration and conflict reduction. Those properties are easily tested at a linguistic level, due to the fuzzy common partitioning.

We consider, in table 1, different types of relations between the input space covered by two rules and two cases of rule consequences: they can be either equal or different. It would be possible to extend this study to consider the inclusion relation in the output space.

In the following, the space covered by a rule is understood at a linguistic level, and not at a numerical one: the rules " $If\ X\ is\ MF_1 \ldots$" and " $If\ X\ is\ MF_2 \ldots$" cover distinct parts of the input space defined by "MF_1" and "$MF2$", even the corresponding membership functions overlap.

Input \ Output	=	$\neq$
=	Redundant	Inconsistent
$\subset$	Redundant	Inconsistent (specification)
$\cap \neq \emptyset$	Partially redundant	Partially inconsistent
$\neq$	No problem	

Table 1: Conflicts to be analyzed from rule description

In table 1 it is obvious that if the spaces covered by the rules are the same, then one could be considered as included in the other and their intersection is not empty. Thus, *included* should be interpreted as different but one included into the other, *not empty intersection* should be interpreted as spaces not included one into the other but with a non empty intersection, and similarly, *different* means with no common part. In addition, the expression *Inconsistent (specification)* means that under some hierarchical approaches (Yager, 1993) this situation could be managed as a specific rule that overwrites a more general rule.

Figure 5 shows an illustration of the *included* case while figure 6 illustrates the *not empty intersection* one.

The detection of this kind of conflicts at the linguistic level is quite easy. The analysis on conflict reduction under the control of the expert is described in (Guillaume & Magdalena, 2003).
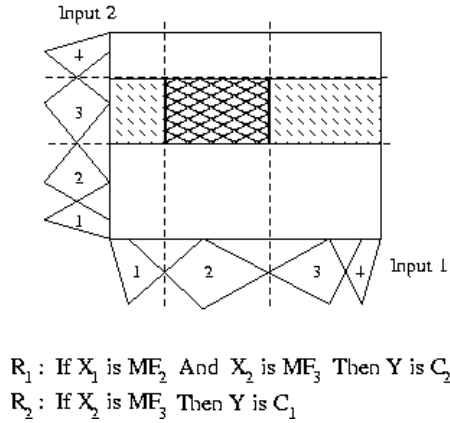
R_1 : If X_1 is MF_2 And X_2 is MF_3 Then Y is C_2
R_2 : If X_2 is MF_3 Then Y is C_1

Figure 5: Input space covered by R1 is included in the one covered by R2

R_1 : If X_1 is MF_2 Then Y is C_2
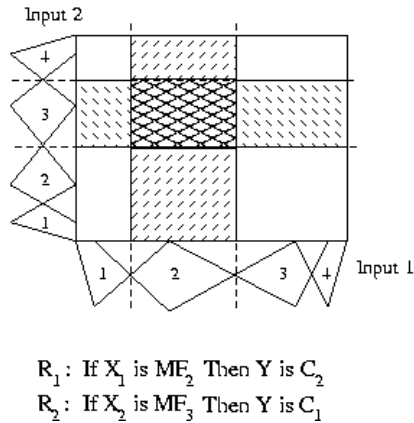R_2 : If X_2 is MF_3 Then Y is C_1

Figure 6: Two rules covering spaces with a non empty intersection

Let us now apply the first step of the whole process, the definition of the common fuzzy partitioning, to real world data.

6 Application example

The present approach is being applied as the methodology of the Knowledge Extraction Tasks of the ADVOCATE II project (`http://advocate2.e-motive.com`). The project concentrates on Intelligent on-board fault detection and recovery in autonomous, underwater and ground, vehicles.

To illustrate some of the previously described ideas a few comments on the knowledge extraction process for an Autonomous Ground Vehicle called BART (Basic Agent for Robotic Tasks) are introduced.

Concretely we briefly describe the process of defining a fuzzy partition for one of the variables related to a specific problem to be diagnosed (Guillaume, Magdalena, Sotelo, Bergasa, Flores & Ocaña, 2003). It is a motion problem related to the collision with an obstacle not detected by sensory means that is dragged by the vehicle. In this situation, an early detection of the problem may reduce system damage.

Once stated the problem, the first expert contribution is the identification of useful variables. The collision may be detected by comparing expected and observed robot motion. The list of variables to be considered, after appropriate preprocessing are:

1. Difference of battery voltage.

2. Commanded and measured linear velocity.

3. Commanded and measured angular velocity.

4. Depth and width of velocity undershoot.

5. Ultrasound range measurements and their derivatives.

To analyze the problem, several experiments have been carried out with different kinds of obstacles and different configurations of the obstacle-robot situation. The objective of these experiments is twofold: to create experience, to be put in the form of expert fuzzy rules, and to generate data to be used for additional knowledge induction.

The experiments have demonstrated that a fast but deep undershoot in vehicle velocity takes place upon collision with an obstacle, until the velocity controller regains the commanded reference. This constitutes the key hint to properly provide a diagnosis on it. Preprocessing through Kalman and FIR (Finite-input response) filtering allows us to define a clearer *velocity undershoot* produced by the collision. But for the experts there was no clear idea of the meaning of this *synthetic variable*. Consequently, for this variable, only a quite general information was provided by the experts while most information in fuzzy partitioning was obtained from experimental data.

Once set up a family of experiments, the obtained undershoots have been used to define the fuzzy partitions. Figures 7 and 8 show fuzzy partitions with different granularity that have been derived by using two partitioning methods implemented in an open source software called FisPro (Guillaume, Charnomordic & Lablée, 2002).
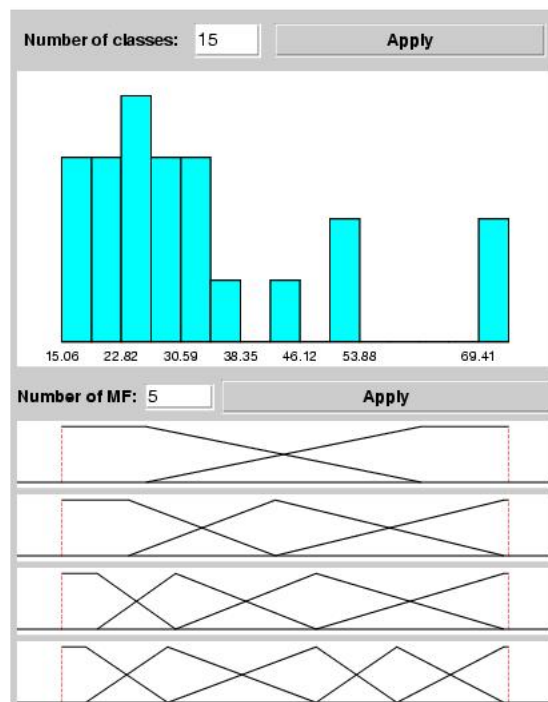

Figure 7: Fuzzy partitions using k-means.

Analyzing the fuzzy partitions, the experts have determined that the best suited result is that obtained by HFP with four fuzzy sets, but adding a fifth term, corresponding to the label *null*, to include the case of no collision that was not represented in the experimental data.

The final result is shown in Figure 9 where data distribution is compared with the selected membership functions. Thanks to the superimposed conditions described in Section 4.1, the partition is highly interpretable. Expert knowledge
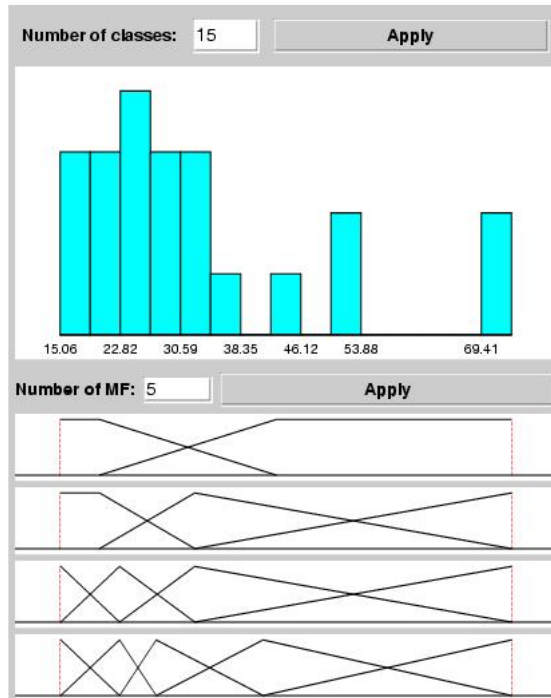
Figure 8: Fuzzy partitions using HFP.

was used to determine the suitable number of terms. Firstly, four induced terms are labeled, from SMALL to HUGE, and then the term NULL is added. On the other hand, data proved useful to quantify the fuzzy sets corresponding to the induced linguistic terms.

7 Conclusion

The cooperation of expert knowledge and data in system design remains an open problem, especially when the goal is to get a system which is both accurate and interpretable.

The two kinds of knowledge, expert knowledge and data, convey complementary information. The objective is to extract their specific contribution to make the cooperation benefit for the system to be designed.

The proposed method described in this paper is made up of two integration levels. In the first one, the data distribution is used to design strong fuzzy partitions for each separate variable under expert control. This type of partitioning ensures each fuzzy set can be attached a linguistic label. Several quality criteria are available to characterize the partitions, but the final semantical agreement is given by the expert: the fuzzy set centers must correspond to possible prototypes of the corresponding labels.

Then, rules defined by these linguistic labels can be written by the expert and induced from data. The second level of integration aims to merge the two rule bases into new one checking the fundamental properties of completeness, consistency and redundancy. The possible effects of consistency and redundancy are analyzed, and, as with the first integration level, the final decision is made by the expert. Completeness is only considered at the integration level as none of the rule base is expected to be complete.

A portable software, which implements the whole process, is currently under design. It will be available soon through the Internet.
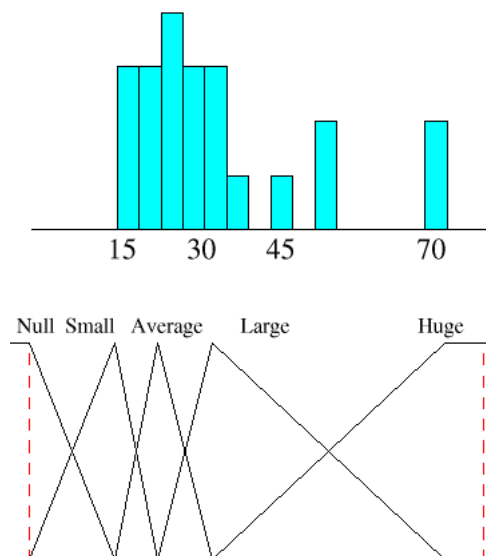
Figure 9: Final interpretable fuzzy partition.

Acknowledgement

References

Bezdek, J. C. (1981). *Pattern Recognition with Fuzzy Objective Functions Algorithms*. Plenum Press, New York.

Bonissone, P. P. (1997). Soft computing: the convergence of emerging reasoning technologies. *Soft Computing*, *1*, 6–18.

Bonissone, P. P., Khedkar, P. S., & Chen, Y. (1996). Genetic algorithms for automated tuning of fuzzy controllers: A transportation application. In *Proc. Fifth IEEE International Conference on Fuzzy Systems (FUZZ-IEEE'96)*, (pp. 674–680)., New Orleans, USA.

Brown, M. & Harris, C. (1994). *Neurofuzzy Adaptive Modelling and Control*. Prentice Hall, Englewood Cliffs.

Casillas, J., Cordon, O., Herrera, F., & Magdalena, L. (2003). *Interpretability Issues in Fuzzy Modeling*, volume 128 of Studies in Fuzziness and Soft Computing. Springer.

Chen, M.-Y. (2002). Establishing interpretable fuzzy models from numerica data. In *Proceedings of the 4th World Congress on Intelligent Control and Automation*, (pp. 1857–1861). IEEE.

Cordón, O. & Herrera, F. (1997). A three-stage evolutionary process for learning descriptive and approximate fuzzy logic controller knowledge bases from examples. *International Journal of Approximate Reasoning*, *17*(4), 369–407.

Cordón, O., Herrera, F., Magdalena, L., & Villar, P. (2001). A genetic learning process for the scaling factors, granularity and contexts of the fuzzy rule-based system data base. *Information Sciences*, *136*(1-4), 85–107.

Cornelissen, A., van der Bergand W.J. Koops, J., & Kaymak, U. (2003). Elicitation of expert knowledge for fuzzy evaluation of agricultural production systems. *Agriculture, Ecosystems and Environment*, *95*, 1–18.

de Oliveira, J. V. (1999). Semantic constraints for membership functions optimization. *IEEE Transactions on Systems, Man and Cybernetics. Part A*, *29*(1), 128–138.

Dunn, J. C. (1973). A fuzzy relative of the isodata process and its use in detecting compact well-separated clusters. *J. Cybernetics*, *3 (3)*, 32–57.

Espinosa, J. & Vandewalle, J. (2000). Constructing fuzzy models with linguistic integrity from numerical data-afreli algorithm. *IEEE Transactions on Fuzzy Systems*, *8 (5)*, 591–600.

Fukuyama, Y. & Sugeno, M. (1989). A new method of choosing the number of clusters for fuzzy c-means method. In *Proceedings of the 5th Fuzzy system symposium (in Japanese)*, (pp. 247–250).

Glorennec, P. Y. (1996). Constrained optimization of FIS using an evolutionary method. In F. Herrera & J. L. Verdegay (Eds.), *Genetic Algorithms and Soft Computing*, number 8 in Studies in Fuzziness and Soft Computing (pp. 349–368). Physica-Verlag.

Glorennec, P.-Y. (1999). *Algorithmes d'apprentissage pour systèmes d'inférence floue*. Editions Hermès, Paris.

Guillaume, S. (2001). Designing fuzzy inference systems from data: an interpretability-oriented review. *IEEE Transactions on Fuzzy Systems*, *9 (3)*, 426–443.

Guillaume, S. & Charnomordic, B. (2003). *A new method for inducing a set of interpretable fuzzy partitions and fuzzy inference systems from data*, (pp. 148–175). Volume 128 of Studies in Fuzziness and Soft Computing of (Casillas, Cordon, Herrera & Magdalena, 2003).

Guillaume, S. & Charnomordic, B. (2004). Generating an interpretable family of fuzzy partitions. *IEEE Transactions on Fuzzy Systems*, *In Press*.

Guillaume, S., Charnomordic, B., & Lablée, J.-L. (2002). Fispro: An open source portable software for fuzzy inference systems. *http://www.inra.fr/bia/M/fispro*.

Guillaume, S. & Magdalena, L. (2003). Typologie des conflits dans les systèmes de mamdani. In *LFA'03*, (pp. 73–80)., France. Cépaduès Editions.

Guillaume, S., Magdalena, L., Sotelo, M. A., Bergasa, L. M., Flores, R., & Ocaña, M. (2003). Expert and induced knowledge for intelligent motion analysis of ground robots. In *EUSFLAT'03*, Zittau, Germany.

Homaifar, A. & McCormick, E. (1995). Simultaneous design of membership functions and rule sets for fuzzy controllers using genetic algorithms. *IEEE Transactions on Fuzzy Systems*, *3 (2)*, 129–139.

Ishibuchi, H. & Murata, T. (1996). *A genetic-algorithm-based fuzzy partition method for pattern classification problems in Genetic algorithms and soft computing, Number 8 in Studies in Fuzziness and Soft Computing*, (pp. 555–578). Physica-Verlag.

Jiménez, F., Gómez-Skarmeta, A., Roubos, H., & Babuška, R. (2001). A multi-objective evolutionary algorithm for fuzzy modeling. In *Proceedings of the 9th IFSA World Congress and the 20th NAFIPS International Conference*, (pp. 1222–1228)., Vancouver, Canada.

Jin, Y., von Seelen, W., & Sendhoff, B. (1999). On generating FC^3 fuzzy rule systems from data using evolutionary strategies. *IEEE Transactions on Systems, Man and Cybernetics - Part B: Cybernetics*, *29*(6), 829–845.

Karr, C. (1991). Genetic algorithms for fuzzy controllers. *AI Expert*, *6*(2), 26–33.

Krishnakumar, K. & Satyadas, A. (1995). GA-optimized fuzzy controller for spacecraft attitude control. In J. Periaux, G. Winter, M. Galán, & P. Cuesta (Eds.), *Genetic Algorithms in Engineering and Computer Science* (pp. 305–320). John Wiley and Sons.

Larichev, O. I. (2002). Close imitation of expert knowledge: the problem and methods. *International Journal of Information Technology & Decision Making*, *1*(1), 27–42.

Liao, T., Celmins, A., & Hammell II, R. J. (2003). A fuzzy c-means variant for the generation of fuzzy term sets. *Fuzzy Sets and Systems*, *135*, 241–257.

López, S., Magdalena, L., & Velasco, J. R. (1999). Genetic fuzzy c-means algorithm for the automatic generation of fuzzy partitions. In B. Bouchon-Meunier, R. Yager, & L. Zadeh (Eds.), *Information, Uncertainty, Fusion* (pp. 407–418). Norwell, MA, USA: Kluwer Scientific.

Magdalena, L. & Monasterio-Huelin, F. (1997). A fuzzy logic controller with learning through the evolution of its knowledge base. *International Journal of Approximate Reasoning*, *16 (3)*, 3335–358.

Mamdani, E. H. & Assilian, S. (1975). An experiment in linguistic synthesis with a fuzzy logic controller. *International journal of Man-Machine Studies*, *7*, 1–13.

Medasani, S., Kim, J., & Krishnapuram, R. (1998). An overview of membership function generation techniques for pattern recognition. *International Journal of Approximate Reasoning*, *19*, 391–417.

Nauck, D. & Kruse, R. (1999). Neuro-fuzzy systems for function approximaton. *Fuzzy Sets and Systems*, *101(2)*, 261–271.

Pedrycz, W. (1993). *Fuzzy control and fuzzy systems* (2nd ed.). Studies in Fuzziness. Research Studies Press Ltd, Taunton, Somerset, England.

Pedrycz, W. (1994). Why triangular membership functions? *Fuzzy sets and Systems*, *64 (1)*, 21–30.

Pedrycz, W. (2001). Fuzzy equalization in the construction of fuzzy sets. *Fuzzy Sets and Systems*, *119(2)*, 329–335.

Pedrycz, W. & Vukovich, G. (2002). On elicitation of membership functions. *IEEE Trans. Systems, Man and Cybernetics-Part A:Systems and humans*, *32 (6)*, 761–767.

Riid, A. & Rüstern, E. (2003). *Transparent Fuzzy Systems in Modelling and Control*, (pp. 452–476). Volume 128 of Studies in Fuzziness and Soft Computing of (Casillas, Cordon, Herrera & Magdalena, 2003).

Roy, A. (2000). On connectionism, rule extraction, and brain-like learning. *IEEE Transactions on Fuzzy Systems*, *8 (2)*, 222–227.

Ruspini, E. H. (1982). *Recent developments in fuzzy clustering.*, (pp. 133–147). Pergamon Press, New York.

Setnes, M., Babuska, R., & Verbruggen, H. (1998). Rule-based modeling: Precision and transparency. *IEEE Trans. Systems, Man and Cybernetics*, *28(1)*, 165–169.

Shreiber, G., Wielinga, B., & Breuker, J. (1993). *KADS- A principled approach to knowledge-based system development*. Knowledge-based systems. Academic Press, Harcourt Brace Jovanovich, Publishers.

Takagi, T. & Sugeno, M. (1985). Fuzzy identification of systems and its applications to modeling and control. *IEEE Transactions on System Man and Cybernetics*, *15*, 116–132.

Turksen, I. B. (1991). Measurement of membership functions and their acquisition. *Fuzzy Sets and Systems*, *40*, 5–38.

Xie, X. & Beni, G. (1991). A validity measure for fuzzy clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *13 (8)*, 841–847.

Yager, R. (1993). On a hierarchical structure for fuzzy modeling and control. *IEEE Trans. on Systems, Man, and Cybernetics*, *23*(4), 1189–1197.

Zadeh, L. A. (1973). Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Transactions on Systems, Man, and Cybernetics*, *3*, 28–44.

Zadeh, L. A. (1975). The concept of a linguistic variable and its applications to approximate reasoning - Parts I, II and III. *Information Sciences*, *8-9*, 199–249, 301–357, 43–80.

Zadeh, L. A. (1997). What is soft computing. *Soft Computing*, *1*, 1.