Appl. Math. and Comp. Sci., 1996, vol.6, No.3, 485–503

AN INTRODUCTION TO FUZZY SYSTEMS

DIDIER DUBOIS*, HENRI PRADE*

This short paper introduces the notion of fuzzy sets as a tool for modelling sets with ill-defined boundaries. Fuzzy sets naturally appear when describing the meaning of natural language words pertaining to quantitative scales, or when modelling the notion of typicality. Three main semantics for fuzzy sets are respectively recalled: similarity, preference and uncertainty. Each semantics underlies a particular class of applications. Similarity notions are exploited in clustering analysis and fuzzy controllers. Uncertainty is captured by fuzzy sets in the framework of possibility theory. The membership function of a fuzzy set is sometimes also a kind of utility function that represents flexible constraints in decision problems.

1. Introduction

The last thirty years have witnessed the growing importance of computer science in the scientific arena, especially engineering, including medical engineering. Nowadays, the computer is not only viewed as an efficient machine for numerical computations, but also as a device potentially capable of storing knowledge in a human-like way, and displaying human-like intelligent behaviour in reasoning and decision tasks. This trend is embodied in recently emerged disciplines such as Artificial Intelligence, Pattern Recognition, Knowledge Engineering in which representation issues are of primary concern.

Among problems faced by knowledge representation is that of mechanizing inference and decision. One reason why this problem is difficult is the existence of several facets to uncertainty, and the discovery that traditional tools for representing uncertainty, such as the error interval analysis and probability theory, are not able to grasp separately all facets of uncertainty. The idea of interval analysis (see e.g. Moore, 1966) is to represent poorly known values by sets, and to engage in computations with functions that have set-valued arguments. However, pieces of subjective knowledge often refer to sets with ill-defined boundaries due to the vagueness pervading terms in natural languages. This state of facts has motivated the introduction of fuzzy sets (Zadeh, 1965).

More particularly, the concept of a fuzzy set deals with the representation of classes whose boundaries are ill-defined, or flexible, by means of characteristic functions taking values in the interval $[0, 1]$. In a fuzzy class, some elements that are considered as "marginal" or "less acceptable", are given a degree of membership that

* Institut de Recherche en Informatique de Toulouse (I.R.I.T.) — C.N.R.S., Université Paul Sabatier, 118 route de Narbonne, 31062 Toulouse Cedex, France, e-mail: {dubois,prade}@irit.fr

is intermediary between 0 and 1. Such fuzzy sets naturally arise in two situations: first, when a finite-term set has to be mapped to a linguistic scale. For instance, if the scale of human sizes is chosen to be the real interval $[0, 250]$ in centimeters, and the term set is $\{$*short, medium-sized, tall*$\}$, then it is difficult to find unquestionable thresholds a and b, such that, e.g. *short* $= [0, a)$, *medium-sized* $= [a, b]$, and *tall* $= (b, 200]$. The predicates *short, medium-sized, tall* are gradual rather than clear-cut. The second situation deals with the representation of typicality, when not all elements of a class are good representatives of that class.

Fuzzy sets (Zadeh, 1965) and possibility theory (Zadeh, 1978; 1987) offer a unified framework for taking into account the gradual or flexible nature of specifications, and the representation of incomplete information. General introductions to fuzzy sets and possibility theory can be found in (Dubois and Prade, 1980; 1988; Dubois *et al.*, 1993; Klir and Folger, 1988; Kruse *et al.*, 1994).

2. Fuzzy Sets

Among early motivations of fuzzy set theory there is a question of representing human-like taxonomies in pattern classification (Bellman *et al.*, 1966). It is sometimes difficult to classify a set of objects into a set of disjoint classes, because the adequacy of an object to the prototypical features that define a class may be a matter of degree. Color is a good example of an attribute that cannot be described by referring to a single class such as *blue, red, yellow, black*, or *white*. Usually colored objects participate to some extent in several of these classes. Hence the boundaries of these classes are gradual rather than clear-cut. Moreover, introducing new intermediate categories like *orange, purple* or *green*, may be useful but does not fix the lack of boundaries; only *more* gradual transitions between categories are thus built. It also leads to a more complex classification structure, hence more difficult to grasp. Fuzzy sets are concerned with flexible data summarization.

Closely related is the question of lexical imprecision in the natural language. When expressing knowledge, individuals would rather use words than numbers, and numbers they might use often refer to orders of magnitude and must not be taken for granted. Understanding a vague term in a natural language assumes that one identifies the universe of concerned objects and the subclass it refers to. Very often, this class has no definite boundaries, and its representation by means of a well-defined set is but a crude, arbitrary approximation. The reason is that predicates that define the class involve gradedness. For instance, "*old* people" is such a class; there is no age beyond which a person is suddenly *old* (even in a given context); he or she *becomes* old. Entering the class is a *continuous* process. The simplest way of modelling gradedness in set-membership, as proposed by Zadeh (1965), is to use a numerical scale, the unit interval as it turned out. So, "*old*" can be roughly captured by a (usually continuous) membership function μ_{old} that assigns a "degree of oldness" to each age in the human age scale, in a given context. This representation, although somewhat arbitrary and leading to new problems (how to measure the "oldness"), is far better than a simple set because, regardless of the precise membership values, it restores the intrinsic continuity of the boundary of the class of "*old ages*". In some applications, such as process control using fuzzy logic (Sugeno, 1985) or classification (Bezdek, 1981), preserving continuity is crucial and explains the success of the model. Of course, not all categories in natural languages can be represented as simply as *old*. There are vague categories for which the universe of concerned objects is itself ambiguous, and the attributes they refer to are ill-defined or controversial (e.g. *beautiful*). The range of applicability of fuzzy set theory excludes these categories but is restricted to cases where the universe is well-defined and can be modelled by means of a numerical scale (age, temperature ...) or an indisputable combination thereof. One may also consider fuzzy subsets of finite universes. In that case, the notion of a membership function may capture the idea of typicality. For instance, *bird* is a fuzzy class because some animal species are more bird-like than others (robins are better than penguins). Hence the class *bird* is naturally equipped with an ordering of typicality that can be encoded in a membership function, although not necessarily a numerical one.

3. The Semantics of Fuzzy Sets

Fuzzy sets seem to be relevant in three classes of applications: classification and data analysis, reasoning under uncertainty, and decision-making problems. These three directions that have been investigated by many researchers actually correspond to and/or exploit three semantics of the membership grades, respectively in terms of similarity, uncertainty and preference. Indeed, considering the degree of membership $\mu_F(u)$ of an element u in a fuzzy set F, defined on a referential U, one can find in the literature three interpretations of this degree:

- **Degree of similarity**: $\mu_F(u)$ is the degree of proximity of u to prototype elements of F. Historically, this is the oldest semantics of membership grades (Bellman *et al.*, 1966). This view is particularly suitable in classification, clustering, regression analysis and the like, where the problem is that of abstraction from a set of data. For instance, one might be interested in classifying a car of known dimensions in the category of "big cars". It is clear that this is a matter of degree. To compute this degree one may choose a prototype of a big car like a Mercedes, and construct a measure of distance between our car under consideration and this Mercedes. One may consider that the membership grade of our car to the fuzzy set of big cars is a decreasing function of this distance.

- **Degree of preference**: F represents a set of more or less preferred objects (or values of a decision variable x) and $\mu_F(u)$ represents an intensity of preference in favour of the object u, or the feasibility of selecting u as a value of x. This view is the one later put forward by Bellman and Zadeh (1970); it has given birth to an abundant literature on fuzzy optimization and decision analysis. Applications pertain to design and planning problems. For instance, one may be interested in buying a big car. Then the membership grade of a given tentative car to the class of big cars reflects our degree of satisfaction with this particular car. Note that here the choice of a car is ours. In other words, the variable whose value is the name of the chosen car is controllable.

- **Degree of uncertainty:** this interpretation was proposed by Zadeh (1978) when he introduced possibility theory. The quantity $\mu_F(u)$ is then the degree of possibility that a parameter x has value u, given that all that is known about it is that "x is F". The possibility can convey an epistemic meaning (F then describes the more or less plausible values of x) or a physical meaning ($\mu_F(u)$ being the degree of ease of having $x = u$). Viewing $\mu_F(u)$ as a degree of uncertainty only refers to the epistemic interpretation. The physical meaning of possibility has more to do with preference and feasibility. For instance, somebody says has just seen a big car. In this situation, the membership grade of a given tentative car to the class of big cars reflects our degree of plausibility that this kind of car is the same as the one seen by the person. When this degree is high, our confidence may still be low that we know which car it is, especially if there are several alternatives. However, if this degree is low, then the car can be rejected as a very implausible candidate. In this case, the choice of the car is not ours: a big car passed by and it is what it is. In other words, the variable whose value is the name of the big car is now uncontrollable.

Note that some semantics are in some sense more basic than others when they can be operationally measured: similarity is evaluated via distance indices, and feasibility may be measured by costs (as in toll sets, Dubois and Prade, 1993a). Preference and plausibibility are derived notions, mainly because of their epistemic nature. Preference may be a function of cost; but it may also reflect similarity with respect to an ideal object and this view is often found in the literature on optimization. Plausibility may also be viewed as similarity with respect to a maximally plausible situation. However, it can also be construed as an upper bound on a frequency or a likelihood function. This is how probability and possibility get connected (Dubois and Prade, 1993b).

Another point is that although fuzzy sets have been introduced with a numerical flavour, a membership function is not necessarily mapped on a set of numbers, but an ordered set such as a complete lattice is enough. A membership function can even be construed as an ordering relation $\geq_F$, attached to a predicate F, where $u \geq_F u'$ means that u is more F than u'. A fuzzy relation R on $U \times U$ can also be viewed as a ternary relation (on U^3), i.e. a collection $\{\geq_u,\ u \in U\}$ of binary relations that are complete preorderings. Then $\mu_R(u, u') \geq \mu_R(u, u'')$ can represent a situation where $u' \geq_u u''$, which reads u' is closer to u than u''. These structures are common in conditional logics of counterfactuals (Lewis, 1973).

However, reducing a fuzzy set to a mere ordering relation between elements prevents one from directly defining fuzzy set-theoretic aggregations. It has been well-known for a long time that there is no simple way of aggregating ordering relations (due to Arrow's impossibility theorem). These difficulties are by-passed by assuming that degrees of membership to fuzzy sets pertaining to unrelated concepts are commensurate. This is done by resorting to a common membership scale (that need not be numerical). This commensurability assumption is often taken for granted and never emphasized in the literature. Yet, it is better to state it clearly in the face of newcomers to the theory, as it sheds light on potential limitations of the fuzzy approach and helps locating it within alternative settings such as measurement theory and multiple-criteria decision analysis.

4. Fuzzy Sets and Similarity

The notion of a set and the notion of identity are indissolubly linked together. A set is defined via conformity with a given clear-cut property, and defines a binary partition of the underlying universe: those elements in the set and those elements out of the set. More generally, a family $\{A_1, A_2, \ldots, A_n\}$ of subsets of a set U generates a partition of U where each class in the partition gathers elements which share the same properties. There is a unique associated equivalence relation R that expresses this partitioning: uRv iff for all i, $u \in A_i$ implies and is implied by $v \in A_i$. Such elements are considered identical since indistinguishable using the family of sets. This scheme has been developed by Pawlak (1991) under the name "rough-set theory". The relation R plays the role of a distance function on U : if uRv, then u and v are close and they are far from each other otherwise.

Analogously, fuzzy sets and similarity are tightly linked together. The above program can be carried out for fuzzy sets as well. This has been done by Valverde (1985) and more recently by Klawonn and Kruse (1993a; 1993b). A family $\{F_1, F_2, \ldots, F_n\}$ of fuzzy subsets of a set U generates a similarity relation S. The degree of similarity $\mu_S(u, v)$ reflects the extent to which u and v belong to all fuzzy sets F_i to the same extent:

$$\mu_S(u,v) = \min_{i=1,n} \min\left(\mu_{F_i}(u) * \!\!\rightarrow \mu_{F_i}(v), \mu_{F_i}(v) * \!\!\rightarrow \mu_{F_i}(u)\right)$$

where $* \rightarrow$ denotes a multiple-valued implication related to a fuzzy conjunction operation $*$ by residuation, i.e.

$$a * \!\!\rightarrow b = \sup\left\{x \in [0,1] \mid a * x \leq b\right\}$$

Noticeable choices of operation $*$ are min, product and the "linear" conjunction $\max(0, a + b - 1)$. Note that $a * \!\!\rightarrow b = 1$ as soon as $a \leq b$, and $1 * \!\!\rightarrow b = b$. Such an implication is in accordance with fuzzy set inclusion: F contains G if and only if $\mu_F \leq \mu_G$. The fuzzy relation S verifies properties that make it a generalization of an equivalence relation:

(i) $\forall u \in U,\ \mu_S(u,u) = 1$ (reflexivity)

(ii) $\forall u \in U,\ \forall v \in U,\ \mu_S(u,v) = \mu_S(v,u)$ (symmetry)

(iii) $\forall u \in U,\ \forall v \in U,\ \forall w \in U,\ \mu_S(u,v) * \mu_S(v,w) \leq \mu_S(u,w)$ (max-$*$ transitivity)

The similarity class $[u]_S$ of an element u is the fuzzy set of elements close to it:

$$\mu_{[u]_S}(v) = \mu_S(u,v),\ \forall v \in U$$

A similarity class F should be extensional, in the sense that "the more u belongs to a fuzzy set F, and the more similar u to v with respect to S, the more v belongs to F", which writes

$$\forall u \in U,\ \forall v \in U,\ \mu_F(u) * \mu_S(u,v) \leq \mu_F(v)$$

Similarity relations are closely related to the idea of distance. In particular, if $*$ is linear, $1-\mu_S$ is a pseudometric, and max-$*$ transitivity is equivalent to the triangle inequality; max-min transitivity, which is stronger, corresponds to ultrametrics. This close connection between fuzzy sets and similarity explains why a significant part of the applications of fuzzy sets deals with data analysis (Kacprzyk and Fedrizzi, 1992), clustering and pattern classification (Bezdek and Pal, 1992).

5. Fuzzy Controllers are Similarity-Driven

Another family of applications of ideas of similarity is the fuzzy controller (Berenji, 1992; Lee, 1990; Pedrycz, 1995; Sugeno, 1985). Many authors have pointed out that fuzzy control techniques perform an interpolation between the conclusions of the fuzzy control rules, on the basis of the degrees of matching of the situation under consideration with the condition parts of the different rules. Interpolative reasoning naturally subsumes some idea of proximity since intermediary decisions are computed from the known controls corresponding to the rules applicable to prototypical situations that are the most similar to the current one. So, it is not surprizing that two main fuzzy control methods, namely Mamdani's and Sugeno's, can receive some justifications, starting with a proper handling of proximity relations.

The basic idea was first proposed by Zadeh (1973) and was implemented for the first time by Professor Mamdani's research team (Mamdani and Assilian, 1975) the following year. The principle of this technique is very simple and makes use of only a fragment of fuzzy set theory. The knowledge incorporated in a fuzzy control system is made of a collection of several parallel rules of the type:

$$\text{if } X_1 \text{ is } A_1 \text{ and } \cdots \text{ and } X_p \text{ is } A_p \text{ then } Y \text{ is } B$$

The X_j's are observable variables of the system to be controlled and therefore are the input variables of the control system, whereas Y is the output of the control system and is obviously an input variable for the system that is controlled. The fuzzy sets A_i are elements of fuzzy partitions of the domains of the variables X_j (for instance "negative large" (NL), "negative medium" (NM), "negative small" (NS), "close to zero" (ZE), "positive small" (PS), "positive medium" (PM), "positive large" (PL)), which leads to rules such as "if X_1 is NS and X_2 is PS, then Y is ZE". For instance, it enables us to encode rules like "if the oxygen percentage is high and the temperature is low, then strongly increase the fuel rate" (under the form "if oxy.percent. is PL and temp. is PS, then increase of fuel is PL").

Given precise observations $x_1, \ldots, x_p$, a fuzzy subset $\widehat{B}$ of values of Y is determined by $\mu_{\widehat{B}}(v) = \max_{i=1,n} \mu_{A_1^{(i)}}(x_1) * \cdots * \mu_{A_p^{(i)}}(x_p) * \mu_{B^{(i)}}(v)$, in Mamdani's method ($*$ = min or product). This fuzzy set is a weighted union of the different fuzzy sets, like $B^{(i)}$ (for rule i) which appear in the conclusion parts of the rules whose condition parts more or less correspond to the current observation. This may be seen as a form of analogical reasoning: the value of the control should be taken from the union of the sets of recommended values for the situations (described in the condition parts of the tool for approximating real functions in the above sense, because they are only capable of bracketing functions in the general case (even if this bracketing may be made very small by sufficiently increasing the number of rules), and they do not lay bare a class of simple membership functions with special interpolating capabilities as with neuro-fuzzy methods. For instance, gradual rules with linear membership functions only capture linear interpolation. Even Sugeno's fuzzy interpolation method has some drawbacks from the point of view of interpolation (Babushka, 1995). The potential of fuzzy rules rather lies in their capability to compute linguistic summaries of numerical descriptions of input-output systems inside a prescribed, user-driven vocabulary. For this type of problem the precision of the approximation is less important than the linguistic interpretability of the fuzzy model, in the spirit of early verbal modelling research as done by Wenstøp (1976), for instance. The bracketting property of gradual rule-based approximations of functions can ensure a consistency between a numerical model and the corresponding verbal model.

6. Fuzziness Versus Uncertainty

While surveying the key-papers (Zadeh, 1987), two periods in the development of the fuzzy set theory can be laid bare. The first one (1965–1975) put emphasis on the notion of a fuzzy class as a class without definite boundaries. The main applications of this idea and of the resulting formal tools were fuzzy clustering and classification, a smooth interface between numerical and symbolic knowledge, and the use of interpolative reasoning as done in fuzzy control. From 1975 on, Zadeh developed the idea of a fuzzy set as an elastic constraint on possible situations, parameter values, etc. (Zadeh, 1978). This research culminated in the late seventies and early eighties by the introduction of possibility theory.

There always remains some confusion in the fuzzy literature on the potential of fuzzy sets for handling uncertainty. This state of confusion can be exemplified in some texts by fuzzy set proponents claiming that probability theory models randomness while fuzzy set models subjective uncertainty (hence ignoring subjective probability). It is also present in the expert systems literature where certainty factors have been confused with membership grades. It also pervades the antifuzzy literature where the truth-functionality of conjunction, disjunction and negation in fuzzy logic is considered as mathematically inconsistent (see (Elkan, 1994) for a recent restatement of this fallacy).

In fact, insofar as fuzzy sets model vague predicates, membership grades model degrees of truth of vague propositions, not degrees of uncertainty. In the scope of knowledge representation, truth is a matter of convention, while uncertainty reflects incomplete or contradictory knowledge. In classical logic the convention is that truth is binary. The fuzzy set theory (and before, multivalued logics) has modified this convention. This shift in convention does not entitle degrees of truth to be interpreted as degrees of uncertainty.

Degrees of uncertainty can be attached to a clear-cut proposition in order to model the fact that it is not known whether this proposition is true or false. Uncertainty is at the meta-level with respect to truth. In classical logic truth is binary (a proposition is true or false) while uncertainty is ternary (a proposition is surely true, surely false, or unknown). In probability theory, truth is usually binary (crisp propositions) while uncertainty takes on all values in the unit interval. In the fuzzy set theory truth is many-valued but there is no uncertainty insofar as the element, the membership grade of which is computed, is precisely located. Fuzzy-truth values (which Zadeh has claimed to be typical of fuzzy logic) are uncertainty distributions that describe partially unknown truth-values.

As pointed out above, it happens that a degree of membership $\mu_F(u)$ is interpreted as a degree of uncertainty instead of a degree of truth. But this degree of uncertainty is not attached to the fuzzy proposition 'X is F' when it is known that $X = u$, but to the non-fuzzy proposition $X = u$, when all that is known is that the value of X is somewhere in the support of F.

From the point of view of uncertainty modelling, probability theory is unable to model ignorance (and partial ignorance) in a natural way. Probability theory, which, in decision-theory, has traditionally been considered a proper tool for representing partial belief, is not fully satisfactory for that purpose. Uniform probability distributions on finite sets better express randomness than ignorance: when you ignore the properties of a die, assume it is fair so as to maximize the uncertainty of outcomes. Ignorance rather means that each event is equally plausible, because there is no available evidence that supports any of them. No probability measure can account for such a state of lacking knowledge. The problem of representing ignorance is at the root of controversies regarding the existence of prior probabilities in Bayesian inference. When a fuzzy set is used in order to represent what is known about the value of a variable, the degree attached to a value only expresses the level of possibility that this value is indeed *the* value of the variable. However, in the case of incomplete information, several values may have a degree of possibility equal to 1. All values may have possibility value 1 in the case of complete ignorance.

7. Possibility Theory and Common-Sense Reasoning

The main contents of Zadeh's message on the representation of uncertainty can be summarized as follows:

- A piece of fuzzy knowledge (as given for instance in natural language) pertaining to a given universe of discourse U (other would say a "set of possible worlds") expresses an elastic constraint on this universe, that can be encoded as a possibility distribution π, i.e. a mapping from U to $[0,1]$ such that $\pi(u) = 0$ means that u is an impossible situation;
- Such pieces of knowledge can be used to address queries, and the corresponding responses can be given in terms of possibility-qualified and certainty-qualified statements.

In the following, subsets of the set U of possible worlds are denoted by A, B, $C, \ldots$ and $\emptyset A$ denotes the complement of A. Here π is a possibility distribution that encodes pieces of imprecise knowledge and x denotes the variable fuzzily restricted by π. Given a simple query of the form "does x belong to A?"

rules) that resemble more or less the observation; this union is then obviously weighted by a combination of the similarity degrees $\alpha_j = \mu_{A_j^{(i)}}(x_j)$, $j = 1, p$. In practice (for a continuous ordered domain), a particular value y_0 of $\widehat{B}$ is chosen (operation usually called "defuzzification") often as the barycentre of the elements of $\widehat{B}$ with respect to their membership degrees (method usually called "gravity centre method" in the literature). In the particular case, where the subsets $B^{(i)}$ appearing in the conclusion parts of rules reduce to precise scalar values $b^{(i)}$, this method leads to compute y_0 as a weighted mean of these precise values. This latter combination rule is used by Sugeno and his group (e.g., Sugeno and Nishida, 1985) and by many others in numerous applications. In the case of triangular membership functions (often used in practice) for the $A_j^{(i)}$'s, it appears that in general a precise observation x_j^o only belongs to two fuzzy sets with a strictly positive degree.

The fuzzy rules for control are worked out either from information given by a human operator (as is done for expert systems), or from recordings of a human operator operating the system. Self-adaptation techniques have been proposed and can be used when necessary to determine the fuzzy partitions of the domains (number of fuzzy subsets and parameters connected to their width). Neural networks are also used for fuzzy control rules acquisition (e.g. Lee, 1991). A fuzzy control unit does the same work as a PID controller since it defines implicitly a numerical function tying the control variable and the observed variables together. The difference between classical and fuzzy control methods lies in the way this control law is found. In the context of classical automatic control, especially optimal control theory, the control law is calculated using a mathematical model of the process, whereas the fuzzy logic approach, consistent with artificial intelligence, suggests that the control law be built starting from the expertise of a human operator. In the case of industrial applications of PID controllers, the philosophy is close to fuzzy logic controllers, since the tuning of the PID parameters is usually done in an ad hoc way. However, only linear control laws can be attained with a PID, while the fuzzy controller may capture non-linear laws, which may explain some observed successes of the fuzzy controller over PID controllers. In fact, any kind of control law can be modelled by the fuzzy controller methodology, provided that this law is expressible in terms of "if ..., then ..." rules, just like in the case of expert systems. However, fuzzy logic diverges from the standard expert system approach by providing an interpolation mechanism from several rules. So doing, although a linguistic rule-based control is modelled, the function simulated by the control part remains continuous just like in classical automatic control. If the "right rules" are exhibited, one gets a possibly non-linear continuous control law where jolts are avoided. In the context of complex processes, it may turn out to be practically simpler to get the knowledge from an expert operator rather than to calculate an optimal control.

The fuzzy control systems after having first been developed and tested in Europe at the end of the 1970s, have been applied on a large-scale in Japanese industry since the first half of the 1980s (Sugeno, 1985). Since then, interest in fuzzy control has still increased in Japan, where many industrial applications have been developed: e.g. automatic guidance control for subway trains, chlorine controller for water purification plants, elevator control systems, automatic focusing systems for cameras; system for the control of shield tunnelling, control of automatic container crane operation systems, control system of the speed of a car, stabilization of the image of a camcorder, automatic selection of a program of a washing machine, etc. See (Bellon *et al.*, 1992) for an overview. Recently, Germany has started extensive research programs especially on fuzzy control, both with public money (in Westfalia, especially) and with industrial support (e.g. SIEMENS).

In all these applications, the same interpolation mechanism provided by the fuzzy control methodology is at work: the current situation encountered by the system resembles more or less two or more prototypical situations for which recommended control actions are known, and a control action which is intermediary between these recommended ones is computed on the basis of the resemblance degrees. For instance, in the simple case of the washing machine, the temperature of the water, the washing time and the quantity of washing product are determined from the weight, the nature and the dirtiness of the clothes, using a rather small set of fuzzy rules which roughly describes how these variables are related.

In fuzzy controllers no uncertainty is involved, only similarity-based reasoning is at work. Klawonn and Kruse (1993a; 1993b) have shown that a set of fuzzy rules can be viewed as a set of crisp rules along with a set of similarity relations and that the extensionality axiom applied to the similarity relation induced on the input-output space justifies Mamdani's method. Alternatively, an interpolation-dedicated fuzzy rule "if X is A, then Y is B" can be understood as "the more X is A the more Y is B". This type of fuzzy rule is called "gradual rule" and the corresponding inference machinery works as follows: if $X = x$ and $\alpha = \mu_A(x)$, then Y lies in the level cut B_α containing the values with membership grades at least equal to α. When two rules are at work, such that $\alpha_1 = \mu_{A_1}(x)$, $\alpha_2 = \mu_{A_2}(x)$, then the conclusion $Y \in (B_1)_{\alpha_1} \cap (B_2)_{\alpha_2}$ lies between the cores of B_1 and B_2, i.e. on ordered universes, an interpolation effect is obtained. It can be proved that Sugeno's fuzzy reasoning method for control can be cast in the framework of gradual rules (Dubois, Grabisch and Prade, 1995). More generally, interpolation is clearly a kind of reasoning based on similarity (rather than uncertainty) and it should be related to current research on similarity logics. Similarity relations and fuzzy interpolation methods should also impact on current research in case-based reasoning.

Neither gradual rules, nor other kinds of well-behaved fuzzy rules, provide a tool of the kind used in the neuro-fuzzy literature for approximating functions (e.g. Mendel, 1995). In the approximation problem, what is looked for is a class of fuzzy rules that are simple (e.g. Gaussian membership functions) and an interpolation method that is altogether powerful (a large class of functions can be approximated with arbitrary precision), easy to compute with, and amenable to efficient learning methods when applied to input-output data. Fuzzy rule-based systems are then viewed as a convenient way of compressing data. It has been proved that fuzzy rules are universal approximators, but such results need arbitrary large sets of rules to be achieved, which does not agree with the idea of data compression. Moreover, the initial ambition of fuzzy sets, that is the interface between natural language and numerical models of systems, has kind of disappeared in this program. For instance, gradual rules are not an efficient where A is a prescribed subset of situations, the response to the query can be obtained by computing the uncertainty induced on A by the pieces of fuzzy knowledge, noticeably (Dubois and Prade, 1988):

- A is consistent with π, with degree $\Pi(A) = \sup_{u \in A} \pi(u)$,
- A is certainly implied by π, with degree $N(A) = 1 - \Pi(\emptyset A) = \inf_{u \notin A} 1 - \pi(u)$.

$\Pi(A)$ qualifies the degree of possibility of A, defined by assuming that, if it is only known that A occurs, then the most plausible situation compatible with A prevails. For instance, the degree of possibility of the event "the light is off" is computed after the degree of possibility of "someone switched off the light" rather than "the bulb is broken", insofar as the latter is very unusual compared to the former. Here, "possible" means "unsurprizing". The basic axiom of possibility measures is

$$\forall A,\ \forall B,\ \Pi(A \cup B) = \max\Big(\Pi(A), \Pi(B)\Big)$$

By convention, $\forall A$, $\Pi(A) \in [0,1]$ and $\Pi(\emptyset) = 0$. $N(A) \geq \alpha > 0$ means that the most plausible situation where A is false is rather impossible, i.e. not possible to a level greater than $1 - \alpha$. Set functions Π and N are respectively called possibility and necessity measures, and are simple ordinal models of graded uncertainty (Dubois and Prade, 1988). Their particular character lies in their ordinal nature, i.e. the unit interval is used only to rank-order the various possible situations in U, in terms of their compatibility with the normal course of things as encoded by the possibility distribution π. The above view contrasts with the probabilistic encoding of knowledge, which relies on an additivity assumption. The latter is justified either on the basis of a combinatorial ("case counting") interpretation, or on a pay-off (betting-behavioural) view of uncertainty. Only ordering is requested in possibility theory.

A systematic assumption in possibility theory is that the actual situation is normal, i.e. it is any u such that $\pi(u)$ is maximal given other known constraints. It justifies the evaluation $\Pi(A)$, and contrasts with the probabilistic evaluation of the likelihood of events. Moreover $N(A) > 0$ means that A holds in all the most normal situations. Since the assumption of normality is always made, $N(A) > 0$ means that A is an accepted belief, i.e. one may act as if A were true. This assumption is always a default one and can be revised if further pieces of evidence contradict it.

Possibility theory is an approach to uncertainty where statements of the form "A is possible" and "A is believed" are clearly distinguished via the two set-functions Π and N respectively. This distinction cannot be made in Bayesian probability where the negation of "A is believed " is always interpreted as "not A is believed", instead of "A is not believed" (the latter does not commit into believing "not A"). This feature of possibility theory is shared by alternative settings such as belief functions and upper-lower probabililies. If viewed as genuinely numerical, function N is indeed a particular case of Shafer belief function and lower probability. However the rules of possibility theory make it extremely simple with respect to its competitors, which also means that it has a more limited expressive power since it is purely ordinal.

Fuzzy sets, as possibility distributions, offer a powerful tool for the modelling of various kinds of common-sense reasoning. Let us review some of them.

Fuzzy Deductive Inference

This type of approximate reasoning has been advocated by Zadeh (1975) in the mid-seventies, as a calculus of fuzzy restrictions. The principles of this approach are as follows: consider a set of statements (in natural language, for instance). Each statement is translated into a possibility distribution that relates the considered variables. Then all possibility distributions are conjunctively combined (using minimum operation) into an overall possibility distribution π. Reasoning consists in projecting π over various subsets of variables of interest. This method is a direct generalization of deductive classical logic reasoning which only considers binary-valued possibility distributions: stating proposition p amounts to considering all situations where p is false as impossible. Using fuzzy propositions induced by natural language sentences, the set of situations where a proposition is not false is a fuzzy set, and deductive inference propagates these membership grades, viewed as degrees of possibility. A typical example of fuzzy deductive inference is the generalized modus ponens whereby, from two statements "X is A^*" and "If X is A then Y is B", a conclusion "Y is B^*" can be computed even if A^* slightly differs from A (Zadeh, 1979).

Reasoning under Uncertainty and Inconsistency

A possibilistic knowledge base K is a set of pairs (p, s) where p is a classical logic formula and s is a lower bound of a degree of necessity $(N(p) \geq s)$. It can be viewed as a stratified deductive data base where the higher s, the safer the piece of knowledge p. Reasoning from K means using the safest part of K to make inference, whenever possible. Setting $K_\alpha = \{p, (p, s) \in K, s \geq \alpha\}$, the entailment $K \vdash (p, \alpha)$ means that $K_\alpha \vdash p$. K can be inconsistent and its inconsistency degree is $\mathrm{inc}(K) = \sup\{\alpha, K \vdash (\bot, \alpha)\}$, where $\bot$ denotes the contradiction. In contrast to classical logic, inference in the presence of inconsistency becomes non-trivial. This is the case when $K \vdash (p, \alpha)$, where $\alpha > \mathrm{inc}(K)$. Then it means that p follows from a consistent and safe part of K (at least at level α). This kind of syntactic non-trivial inference is sound and complete with respect to the non-monotonic preferential entailment defined below (Dubois, Lang and Prade, 1994). Moreover, adding p to K and nontrivially entailing q from $K \cup \{p\}$ corresponds to revising K upon learning p, and having q as a consequence of the revised knowledge base. This notion of revision is exactly the one studied by Gärdenfors (1988) at the axiomatic level.

Nonmonotonic Plausible Inference Using Generic Knowledge

Possibilistic logic does not allow for a direct encoding of pieces of generic knowledge such as "birds fly". However, it provides a target language in which plausible inference from generic knowledge can be achieved in the face of incomplete evidence. In possibility theory, "p generally entails q" is understood as "$p \wedge q$ is a more plausible situation than $p \wedge \neg q$". It defines a constraint of the form $\Pi(p \wedge q) > \Pi(p \wedge \neg q)$ that restricts the set of possibility distributions. Given a set S of generic knowledge statements of the form "p_i generally entails q_i", a possibilistic base can be computed as follows. For each interpretation ω of the language, the maximal possibility degree $\pi(\omega)$ is computed, that obeys the set of contraints in S. This is done by virtue of the principle of minimal specificity (or commitment) that assumes each situation as a possible one insofar as it has not been ruled out. Then each generic statement is turned into a material implication $\neg p_i \vee q_i$, to which the level $N(\neg p_i \vee q_i)$ is attached. It comes down, as shown in (Benferhat *et al.*, 1992) to rank-order the generic rules giving priority to the most specific ones. A very important property of this approach is that it is exception-tolerant. It offers a convenient framework for implementing a basic form of nonmonotonic system called "rational closure" (Lehmann and Magidor, 1992), and addresses a basic problem in the expert system literature, i.e. handling exceptions in uncertain rules.

Hypothetical Reasoning

The idea is to cope with incomplete information by explicitly handling assumptions under which conclusions can be derived. To this end some literals in the language are distinguished as being asssumptions. Possibilistic logic offers a tool for reasoning with assumptions. It is based on the fact that in possibilistic logic the clause $(\neg h \vee q, \alpha)$ is semantically equivalent to the formula q with symbolic weight $\min(a, t(h))$, where $t(h)$ is the (possibly unknown) truth value of h. The set of environments in which a proposition p is true can thus be calculated by putting all assumptions in the weight slots, carrying out possibilistic inference so as to derive p. The subsets of assumptions under which p is true with more or less certainty can be retrieved from the weight attached to p. This technique can be used to detect minimal inconsistent subsets of a propositional knowledge base (see Benferhat *et al.*, 1994).

Abductive Reasoning

Abductive reasoning is viewed as the task of retrieving plausible explanations of available observations on the basis of causal knowledge. In fuzzy-set theory causal knowledge has often been represented by means of fuzzy relations relating a set of causes C to a set of observations S. However the problem of the semantics of this relation has often been overlooked. Here $\mu_R(c,s)$ may be viewed either as a degree of intensity or a degree of uncertainty. Namely, when observations are not binary, $\mu_R(c,s)$ can be understood as the intensity of presence of the observed symptom s when the cause c is present. This is the traditional view in fuzzy-set theory. It leads to Sanchez's approach (1977) to abduction, based on fuzzy relational equations. Another view has been recently proposed, where $\mu_R(c,s)$ is understood as the degree of certainty that a binary symptom s is present when c is present. A dual causal matrix R' must be used, where $\mu_{R'}(c,s)$ is the degree of certainty that a binary symptom s is absent when c is present. On such a basis the theory of parsimonious covering for causal diagnosis by Peng and Reggia (1990) can be extended to the case of uncertain causal knowledge and incomplete observations. This method is currently applied to satellite failure diagnosis (Cayrac *et al.*, 1994).

8. Flexible Constraint Propagation

As it turns out, there are two distinct understandings of a possibility distribution. A possibility distribution may encode imprecise knowledge about a situation. In that case, $\pi(u) = 1$ means that u is a normal (maximally plausible) situation. However, in that case, no choice is at stake, that is, the actual situation is what it is and π encodes plausible guesses about it. The other understanding is in terms of preference and leads to a calculus of flexible constraints (Dubois, Fargier and Prade, 1994). Then a possibility distribution encodes a flexible requirement, $\pi(u) = 1$ means that u is a preferred choice, and $\pi(u) > \pi(u')$ means that u is a better choice than u'.

The notion of a constraint is basic in operations research. A constraint describes what are the potentially acceptable decisions (the solutions to a problem) and what are the absolutely unacceptable ones: it is an all-or-nothing matter. Moreover, no constraint can be violated, i.e. a constraint is classically considered as imperative. Especially the violation of a constraint cannot be compensated by the satisfaction of another one. If a solution violates a single constraint, it is regarded as unfeasible. The idea of a goal is quite different: attaining an objective is a matter of degree and leads to introducing a preference ordering on the the solutions, via the objective function attached to a goal; moreover, in the presence of multiple conflicting goals, trade-offs are allowed. In other words, the "optimal" solutions may fail to attain some of the goals, so that goals are not imperative, strictly speaking. In this picture, flexible constraints fall between these two modelling attitudes. The idea is to keep the non-compensatory property of constraints, while introducing intermediary levels between feasibility and non-feasibility, as well as levels in the imperativeness of constraints.

A classical hard constraint C is represented by a classical set of solutions, i.e. only using degrees of membership 0 or 1. However, since eventually a single solution will be picked up, the feasible solutions are mutually exclusive and the characteristic function μ_C that is attached to the constraint is a binary possibility distribution π_x, where x is a vector of decision variables. Introducing intermediary levels of feasibility, C is called a fuzzy, or soft constraint: $\mu_C(s) = 1$ means that a solution s totally satisfies C while $\mu_C(s) = 0$ means that it totally violates C (s is unfeasible). If $0 < \mu_C(s) < 1$, s satisfies C only partially; $\mu_C(s) < \mu_C(s')$ indicates that C is more satisfied by s than by s' (C prefers s to s'). Hence, like an objective function, a fuzzy constraint rank-orders the feasible decisions. However, contrary to an objective function, a fuzzy constraint also models a threshold (represented by the bottom level 0) beyond which a solution will be rejected. In fact, a fuzzy constraint can be viewed as the association of a constraint (defining the support of C) and a criterion which rank-orders the solutions satisfying the constraints. In this interpretive framework, a membership function, construed as a possibility distribution, is similar to a utility function, or better, a value function. A soft constraint C will be looser than another C' if and only if $\mu_C \leq \mu_{C'}$, that is, if any solution to C is at least as feasible as for C'.

Here, the meaning of possibility has to do with feasibility (objective interpretation) or preference (subjective interpretation). The axiom of possibility measures means that if A or B is to be achieved, it is equivalent to achieve the easiest of the two. It has little to do with uncertainty, since all the variables involved in the problem are supposed to be decision variables, i.e. the ultimate choice of the solution is ours. Possibility distributions encode feasibility profiles of soft constraints. In the presence of very loose constraints, it is clear that the use of soft constraints can help breaking ties among feasible solutions, just as objective functions do. On the contrary, when hard constraints are tight, there is no feasible solution. One way out is to relax the constraints, but automating this process is not so easy and is usually time-consuming. What is usually done is to assign priorities to constraints. Solutions that satisfy all constraints are selected if any. Otherwise the less prioritary constraints are dropped, and the chosen solution satisfies only higher priority constraints. It can be proved that this decision strategy can be captured in the setting of possibility theory.

Assume the unit interval is viewed as a priority scale with top 1 and bottom 0. A prioritized constraint is a pair (C, p) where p estimates how imperative C is. $(C, 1)$ denotes an imperative constraint; $(C, 0)$ denotes a "non-constraint", i.e. C can be violated in any case. Priorities on constraints can be transformed, without any loss of information, into feasibility levels on solutions. Indeed, since p represents to what extent it is necessary to satisfy C, $1 - p$ indicates to what extent it is possible to violate it. In other words, any potential solution s that violates C satisfies (C, p) to a degree equal to $1 - p$. While possibility degrees are interpreted in terms of feasibility, necessity degrees are levels of priority. The pair (C, p), where C is a crisp constraint, can be modelled as a special kind of fuzzy constraint C' (Dubois, Fargier and Prade, 1994):

$$\mu_{C'}(s) = \begin{cases} 1 & \text{if } s \text{ satisfies } C \\ 1 - p & \text{if } s \text{ violates } C \end{cases}$$

If C is itself a soft constraint with priority p, it can be modelled by the fuzzy constraint C' with membership function: $\mu_{C'}(s) = \max(1 - p, \mu_C(s))$. The notion of fuzzy set as viewed by Bellman and Zadeh (1970) meant to represent constraints as well as objective functions by fuzzy subsets C of possible decisions. If $C_1, \ldots, C_n$ denote n fuzzy constraints, the fuzzy decision set was defined by

$$\mu_D(s) = \min_{i=1,n} \mu_{C_i}(s)$$

However this definition of a decision set is not really in accordance with the usual view of multiple-criteria decision-making since an optimal solution in the sense of Bellman and Zadeh does not make a trade-off between the membership values $\mu_{C_i}(s)$. On the contrary, an optimal solution is one that least violates the most violated constraint. Hence, Bellman and Zadeh's proposal is a constraint-directed view of problem-solving, where constraint are flexible: if there is a solution s that completely satisfies all constraints, this solution is optimal. Otherwise $\lambda * = \sup_s \mu_D(s) < 1$, and this indicates that the constraints are partially contradictory. To accept a solution such that $\mu_D(s) = \lambda *$ means to partially relax some of the constraints C_i. In that sense constraints C_i are flexible, since the formulation of the problem makes it possible to accept a solution as long as $\mu_D(s) > 0$, which implies some relaxed constraints whenever $\mu_D(s) < 1$.

This notion of optimality is not new in the literature of decision making. It is known as the maximin stategy and is well-known in game theory as the most cautious strategy for players. It is also well-known in decision-making under uncertainty, where there are n states of the world and $\mu_{C_i}(s)$ is the reward of decision s when the state is i; again it is a cautious strategy. Despite these formal analogies, the problem tackled by flexible constraint satisfaction is very different, since index i pertains neither to an opposing player nor to a state of the world. The paradigm most akin to Bellman and Zadeh's is to be found in social choice theory (Moulin, 1988) and $\min_{i=1,n} \mu_{C_i}(s)$ is an egalitarist social welfare function, $\mu_{C_i}(s)$ being the welfare index of an individual i. Maximizing $\min_{i=1,n} \mu_{C_i}(s)$ tends to select solutions which equally satisfy all constraints.

A powerful advantage of the flexible constraint satisfaction setting is its ability to extend all the results and tools from constraint propagation in Artificial Intelligence. Here, preference propagation can be achieved. It is formally equivalent to fuzzy deductive inference, in the sense of Zadeh (1979). Fuzzy inference comes down to consistency analysis (such as arc-consistency, path-consistency, etc.) in the terminology of constraint-directed reasoning (see Dubois, Fargier and Prade, 1994). Fuzzy constraint propagation is carried out to speed up the search for an optimal solution to a problem involving flexible constraints and priorities. Typical examples of application of fuzzy constraint satisfaction are engineering design (Wood *et al.*, 1992) and scheduling (Dubois, Fargier and Prade, 1995).

9. Conclusion

This paper has emphasized the capabilities of fuzzy set-based methods for modelling gradual properties, similarity, flexible constraints, preferences as well as for representing uncertain and imprecise pieces of information. Fuzzy controllers which have become the most popular and visible side of fuzzy set theory are only the emerged part of the fuzzy iceberg. The idea of similarity takes an important place in everyday common-sense reasoning patterns. Proximity and similarity are notions which are naturally graded, which suggests fuzzy sets as a good framework for representing them. Qualitative reasoning, interpolative reasoning, case-based and analogical reasoning can benefit from the proper handling of fuzzy similarity relations. They can also be useful in advanced information systems (e.g., Bosc and Kacprzyk, 1995). Possibility theory offers a convenient framework for handling dual notions like certainty/possibility in the case of incomplete information, or like satisfaction/priority in decision processes. These two features may appear alone or jointly in a great number of fields of applications different from fuzzy control, especially in

- decision support systems, for handling multicriteria aggregation, possibly in imprecise or uncertain environment (Dubois and Prade, 1988; Kaufmann and Gupta, 1988; Zimmermann, 1987);
- advanced information systems, in order to allow for flexible queries and imprecise or uncertain pieces of information stored in the data base (Dubois, Prade and Sanchez, 1989; Bosc and Kacprzyk, 1995);
- the qualitative modelling of complex systems (for instance with human components) (Kickert, 1978), by means of equations involving fuzzy numbers or fuzzy functions, by means of fuzzy "if ..., then ..." rules, or by means of fuzzy regression analysis techniques (Kacprzyk and Fedrizzi, 1992);

- operations research by offering a tool for dealing with flexible constraints, prioritized constraints or constraints with ill-known coefficients for example (Kaufmann and Gupta, 1988; Slowinski and Teghem, 1990; Zimmermann, 1987);
- classification or clustering analysis, by allowing for classes with unsharp boundaries if necessary (when clusters of data are ill-separated) (Bezdek and Pal, 1992).

References

Babushka R. (1995): *Fuzzy modelling: A control engineering perspective.* — Proc. Int. Joint Conf. of *4th IEEE Int. Conf. Fuzzy Systems (FUZZ-IEEE'95)* and 2nd Int. *Fuzzy Engineering* Symp. (IFES'95), Yokohama, Japan, pp.1897-1902.

Bellman R., Kalaba L. and Zadeh L.A. (1966): *Abstraction and pattern classification.* — J. Mathematical Analysis and Applications, v.13, pp.1–7.

Bellman R. and Zadeh L.A. (1970): *Decision making in a fuzzy environment.* — Management Science, v.17, B141–B164.

Bellon C., Bosc P. and Prade H. (1992): *Fuzzy boom in Japan.* — Int. J. Intelligent Systems, v.7, No.4, pp.293–316.

Benferhat S., Dubois D., Lang J. and Prade H. (1994): *Hypothetical reasoning in possibilistic logic: Basic notions, applications and implementation issues*, In: Between Mind and Computer, Fuzzy Science and Engineering, v.1 (Wang P.Z., Loe K.F., Eds.). — Singapore: World Scientific Publ., pp.1–29.

Benferhat S., Dubois D. and Prade H. (1992): *Representing default rules in possibilistic logic.* — Proc. 3rd Int. Conf. *Principles of Knowledge Representation and Reasoning (KR'92)*, Cambridge, MA, Oct.26–29, pp.673–684.

Berenji H.R. (1992): *Fuzzy logic controllers*, In: An Introduction to Fuzzy Logic Applications in Intelligent Systems (Yager R.R. and Zadeh L.A., Eds.). — Boston: Kluwer Academic Publ., pp.68–91.

Bezdek J. (1981): *Pattern Recognition with Fuzzy Objective Function Algorithms.* — New York: Plenum Press.

Bezdek J. and Pal S.K. (1992): *Fuzzy Models for Pattern Recognition.* — New York: IEEE Press.

Bosc P. and Kacprzyk J., Eds. (1995): *Fuzziness in Database Management Systems.* — Heidelberg: Physica-Verlag.

Cayrac D., Dubois D., Haziza M. and Prade H. (1994): *Possibility theory in "fault mode effect analyses" – A satellite fault diagnosis application.* — Proc. 3rd IEEE Int. Conf. *Fuzzy Systems (FUZZ-IEEE 94)*, Orlando, FL, June 26–29, pp.1176–1181.

Dubois D., Fargier H. and Prade H. (1994): *Propagation and satisfaction of flexible constraints*, In: Fuzzy Sets, Neural Networks and Soft Computing (Yager R.R. and Zadeh L.A., Eds.). — New York: Van Nostrand Reinhold, pp.166-187.

Dubois D., Fargier H. and Prade H. (1995): *Fuzzy constraints in job-shop scheduling.* — J. Intelligent Manufacturing, v.6, pp.215–234.

Dubois D., Grabisch M. and Prade H. (1995): *Gradual rules and the approximation of control laws*, In: Fuzzy Rules, Theoretical Aspects of Fuzzy Control (Nguyen H.T., Sugeno M., Tong R. and Yager R., Eds.). — New York: Wiley, pp.147–181.

Dubois D., Lang J. and Prade H. (1994): *Possibilistic logic*, In: Handbook of Logic in Artificial Intelligence and Logic Programming, v.3 (Gabbay D.M., Hogger C.J., Robinson J.A. and Nute D., Eds.). — Oxford: Oxford University Press, pp.439–513.

Dubois D. and Prade H. (1980): *Fuzzy Sets and Systems: Theory and Applications.* — New York: Academic Press.

Dubois, D. and Prade H. (1988): *Possibility Theory — An Approach to the Computerized Processing of Uncertainty.* — New York: Plenum Press.

Dubois D. and Prade H. (1993a): *Toll sets and toll logic*, In: Fuzzy Logic: State of the Art (Lowen R. and Roubens M., Eds.). — Dordrecht: Kluwer Academic Publ., pp.169–177.

Dubois D. and Prade H. (1993b): *Fuzzy sets and probability: Misunderstandings, bridges and gaps.* — Proc. 2nd IEEE Int. Conf. *Fuzzy Systems (FUZZ-IEEE'93)*, San Francisco, CA, March 28–April 1st, pp.1059–1068.

Dubois D., Prade H. and Sanchez E., Eds. (1989): *Editorial.* — Special Issue of Information Systems on Fuzzy Data Bases, v.14, No.6, pp.vii–viii.

Dubois D., Prade H. and Yager R.R., Eds. (1993): *Readings in Fuzzy Sets for Intelligent Systems.* — San Mateo, CA: Morgan & Kaufmann.

Elkan C. (1994): *The paradoxical success of fuzzy logic* (with replies). — IEEE Expert, v.9, No.4, pp.3–49.

Gärdenfors P. (1988): *Knowledge in Flux.* — New York: The MIT Press.

Kacprzyk J. and Fedrizzi M., Eds. (1992): *Fuzzy Regression Analysis.* — Warsaw: Omnitech Press.

Kaufmann A. and Gupta M.M. (1988): *Fuzzy Mathematical Models in Engineering and Management Sciences.* — Amsterdam: North-Holland.

Kickert W.J.M. (1978): *Fuzzy Theories on Decision Making.* — Frontiers in Systems Research, v.3, Leiden: Martinus Nijhoff.

Klawonn F. and Kruse R. (1993a): *Equality relations as a basis for fuzzy control.* — Fuzzy Sets and Systems, v.54, No.2, pp.147–156.

Klawonn F. and Kruse R. (1993b): *Fuzzy control as interpolation on the basis of equality relations.* — Proc. 2nd IEEE Int. Conf. *Fuzzy Systems (FUZZ-IEEE'93)*, San Francisco, CA, March 28–April 1st, pp.1125–1130.

Klir G.J. and Folger T.A. (1988): *Fuzzy Sets, Uncertainty and Information.* — Englewood Cliffs, N.J.: Prentice Hall.

Kruse R., Gebhardt J. and Klawonn F. (1994): *Foundations of Fuzzy Systems.* — New York: Wiley.

Lee C.C. (1990): *Fuzzy logic in control systems: Fuzzy logic controller – Parts 1 and 2.* — IEEE Trans. Systems, Man and Cybernetics, v.20, No.2, pp.404–435.

Lee C.C. (1991): *A self-learning rule-based controller employing approximate reasoning and neural net concepts.* — Int. J. Intelligent Systems, v.6, No.1, pp.71–93.

Lehmann D. and Magidor M. (1992): *What does a conditionnal knowledge base entail?* — Artificial Intelligence, v.55, pp.1–60.

Lewis D. (1973): *Counterfactuals.* — Oxford: Basil Blackwell.

Mamdani E.H. and Assilian S. (1975): *An experiment in linguistic synthesis with a fuzzy logic controller.* — Int. J. Man-Machine Studies, v.7, pp.1–13.

Mendel J. (1995): *Fuzzy logic systems for engineering: A tutorial.* — Proc. IEEE, v.83, No.3, pp.345–377.

Moore R. (1966): *Interval Analysis.* — Englewood Cliffs, N.J.: Prentice-Hall.

Moulin H. (1988): *Axioms of Cooperative Decision Making.* — Cambridge, MA: Cambridge University Press.

Pawlak Z.(1991): *Rough Sets.* — Dordrecht: Kluwer Academic.

Pedrycz W. (1995): *Fuzzy Sets Engineering.* — Boca Raton, FL: CRC Press.

Peng Y. and Reggia (1990): *Abductive Inference Models for Diagnostic Problem-Solving.* — New York: Springer Verlag.

Sanchez E. (1977): *Solutions in conposite fuzzy relations equations — Application to medical diagnosis in Brouwerian logic*, In: Fuzzy Automated and Decision Processes (Gupta M.M. *et al.*, Eds.). — Amsterdam: North-Holland, pp.221–234.

Slowinski R. and Teghem J., Eds. (1990): *Stochastic Versus Fuzzy Approaches to Multiobjective Mathematical Programming under Uncertainty.* — Dordrecht: Kluwer Academic Publ.

Sugeno M. (Ed.) (1985): *Industrial Applications of Fuzzy Control.* — Amsterdam: North-Holland.

Sugeno M. and Nishida M. (1985): *Fuzzy control of model car.* — Fuzzy Sets and Systems, v.16, No.2, pp.103–113.

Valverde L. (1985): *On the structure of F-indistinguishability operators.* — Fuzzy Sets and Systems, v.17, No.3, pp.313–328.

Wenstøp F. (1976): *Fuzzy sets simulation models in a systems dynamics perspective.* — Kybernetes, v.6, pp.209–218.

Wood K.L., Otto K.N. and Antonsson E.K. (1992): *Engineering design calculations with fuzzy parameters.* — Fuzzy Sets and Systems, v.52, No.1, pp.1–20.

Zadeh L.A. (1965): *Fuzzy sets.* — Information and Control, v.8, pp.338–353.

Zadeh L.A. (1973): *Outline of a new approach to the analysis of complex systems and decision processes.* — IEEE Trans. Syst., Man and Cybern., v.3, No.1, pp.28–44.

Zadeh L.A. (1975): *The concept of a linguistic variable and its applications to approximate reasoning — Part I.* — Information Sciences, v.8, pp.301–357.

Zadeh L.A. (1978): *Fuzzy sets as a basis for a theory of possibility.* — Fuzzy Sets and Systems, v.1, No.1, pp.3–28.

Zadeh L.A. (1979): *A theory of approximate reasoning*, In: Machine Intelligence, v.9 (Hayes J.E., Michie D. and Mikulich L.I., Eds.). — New York: Elsevier, pp.149–194.

Zadeh L.A. (1987): *Selected Papers on Fuzzy Sets and Applications*, by Yager R.R., Ovchinnikov S., Tong R.M., Nguyen H.T. (Eds.). — New York: Wiley.

Zimmermann H.J. (1987): *Fuzzy Sets, Decision Making, and Expert Systems.* — Boston: Kluwer Academic Publ.