Chapter 16
Fuzzy Logic in Computer Science

Radim Belohlavek, Rudolf Kruse, and Christian Moewes

What Is Fuzzy Logic?

Motivation

To understand fuzzy logic, it is essential to recall the basic motivation that led to its emergence. This motivation, articulated in various forms in the early papers on fuzzy logic by Zadeh (1965, 1973), can briefly be described as follows. Classical logic is appropriate for a formalization of reasoning that involves bivalent propositions such as "5 is a prime number", "age of Jan is 9", or "if x is a positive integer and $y = x + 1$ then y is a positive integer", i.e., propositions which may in principle be true or false. In a similar way, classical sets are appropriate for representing collections (of objects) that have sharp, clear-cut boundaries, such as "the collection of all prime numbers less than 100" or "the collection of all U.S. Senators as of September 1, 2010". For any such collection, an arbitrary given object either is or is not a member of it.

Most propositions which people use to communicate information about the outer world are not bivalent. Such propositions are true to a certain degree, rather than being true or false only. As an example, "it is hot outside" is a proposition whose truth depends on the outside temperature. According to our intuition, the higher the temperature, the truer the proposition. To require that the proposition be bivalent

R. Belohlavek (✉)
Palacky University, Olomouc, Czech Republic
e-mail: radim.belohlavek@acm.org

R. Kruse • C. Moewes
Otto-von-Guericke University, Magdeburg, Germany
e-mail: kruse@iws.cs.uni-magdeburg.de; cmoewes@ovgu.de

E.K. Blum and A.V. Aho (eds.), *Computer Science: The Hardware, Software and Heart of It*, DOI 10.1007/978-1-4614-1168-0_16,

means to require the existence of a particular value, t, such that the proposition is true if the actual temperature is larger than or equal to t and false if the actual temperature is smaller than t. This means that if the actual temperature is, say, $t - 0.01$, we consider the proposition false, while if it is $t + 0.01$, we consider the proposition true. Therefore, if the proposition "it is hot" is regarded as bivalent, an arbitrarily small change in the outside temperature can change its truth value from false to true and vice versa. Needless to say, this contradicts our intuition and the way we use propositions such as "it is hot outside".

Likewise, most collections of objects to which people refer when communicating information do not have sharp, clear-cut boundaries. The membership of objects in such collections is a matter of degree, rather than being a member or not being a member only. The point is well illustrated by a quote from Zadeh's seminal paper (Zadeh 1965):

> More often than not, the classes of objects encountered in the real physical world do not have precisely defined criteria of membership. For example, the class of animals clearly includes dogs, horses, birds, etc. as its members, and clearly excludes objects as rocks, fluids, plants, etc. However, such objects as starfish, bacteria, etc. have an ambiguous status with respect to the class of animals. The same kind of ambiguity arises in the case of a number such as 10 in relation to the "class" of all real numbers which are much greater than 1.
>
> Clearly, the "class of all real numbers that are much greater than 1," or "the class of beautiful women," or "the class of tall men" do not constitute classes or sets in the usual mathematical sense of these terms. Yet, the fact remains that such imprecisely defined "classes" play an important role in human thinking ...
>
> The purpose of this note is to explore in a preliminary way some of the basic properties and implications of a concept which may be of use in dealing with "classes" of the type cited above. The concept in question is that of a fuzzy set, that is a "class" with a continuum of grades of membership.

Since most propositions about the outer world are not bivalent, classical logic is inadequate to formalize reasoning that involves such propositions. Likewise, since most collections referred to in human communication do not have sharp boundaries, classical sets are inadequate to represent such collections. The main aim of fuzzy logic is to overcome the above-described inadequacies of classical logic and classical sets.

Graded Approach

The principal idea employed by fuzzy logic is to allow for a partially ordered scale of truth values, called also *truth degrees*, which contains the values representing false and true but possibly also other, intermediary truth degrees. That is, the two-element set $\{0, 1\}$ of truth values of classical logic, where 0 and 1 represent false and true, respectively, is replaced in fuzzy logic by a partially ordered scale of truth degrees with the smallest degree being 0 and the largest one being 1. This is known as the *graded approach*. An important example of such scale is the interval [0, 1] of real numbers. A degree from a given scale (e.g., the number 0.9 from [0, 1]) that is assigned to a proposition is interpreted as the degree to which the proposition is considered true. For the proposition "it is hot outside", the higher the outside temperature, the higher the truth degree assigned to this proposition. If 0.9 is assigned to this proposition, it indicates that we consider it being almost hot outside but not completely hot. On the other hand, assigning 0.3 to the same proposition indicates that we consider it being somewhat warm outside but not much. In a similar spirit, scales of truth degrees are used in fuzzy sets to represent degrees to which a given object is a member of a collection with non-sharp boundary. For example, if 0.8 and 0.9 represent degrees to which John and Paul are members of the collection of tall men, respectively, it indicates that both are considered almost tall and that Paul is a little bit taller than John.

Controversies

It is clear from the discussion above that fuzzy logic departs from two important traditions of science – the principle of bivalence and the principle that all scientifically relevant concepts are precise and clear-cut. This departure brought up several fundamental issues at stake, which have been, and continue to be, an object of controversy. Two such issues are briefly described in this section. Another one is discussed in section "Fuzzy Logic and Probability".

The basic idea of fuzzy logic, i.e., that propositions may have intermediary truth degrees, represents a radical departure from one of the basic principles of classical logic and exact sciences – the *principle of bivalence*, according to which every proposition is either true or false. Various ramifications of admitting intermediary truth degrees have been examined in a number of papers, see Smith (2009) for numerous references. Some of the papers pose interesting problems and challenges for fuzzy logic. Quite often, however, the authors of the critical papers are not familiar enough with the principles of fuzzy logic and their analyzes are based on various types of misunderstanding and misconception. Among the critiques of fuzzy logic is a number of attempts to prove that fuzzy logic leads to counterintuitive results and even to contradictions. The best known such critique are Elkan's papers (Elkan 1993, 1994), the second of which appeared in a special issue of *IEEE Expert* along with responses to it. The central claim of Elkan's critique was that "as a formal system, a standard version of fuzzy logic collapses mathematically to two-valued logic." This claim is the content of two theorems presented in Elkan (1993, 1994). In both cases, proofs of the theorems are quite long. Since it is common to take the length of a proof as a measure of profundity of the proven theorem, Elkan's theorems may look on the surface as quite profound. However, a close examination of the theorems demonstrates the contrary. Namely, Belohlavek and Klir (2007) present short proofs of both theorems and by using these proofs they show that axioms upon which Elkan's theorems are based define formal systems that are strange to fuzzy logic and are not capable of dealing with fuzziness.

The second controversy relates to a long-standing tradition in science according to which *all scientifically relevant concepts are precise and clear-cut*. Contrary to this tradition, fuzzy logic claims to provide us with a mathematical tools to model and process concepts that are not clear-cut. Namely, fuzzy logic uses scales of truth values to capture the meaning of propositions and collections which involve non-clear-cut concepts such as “hot”, “tall”, and the like. To capture the meaning of such terms, referred to as *vague terms*, in an appropriate way is quite an intricate issue. This brings up an important question whether the approach of fuzzy logic, based on scales of truth degrees, is appropriate. Such question is very complex and has many facets, ranging from philosophy and mathematics to psychology and cognitive science. Thus far, this question has not been decisively answered and is currently a subject of discussion (van Deemter 2010; Smith 2009). Nevertheless, the use of fuzzy logic is supported by at least the following three arguments. First, fuzzy logic is rooted in the intuitively appealing idea that the truth of propositions used by humans is a matter of degree. An important consequence is that the basic principles and concepts of fuzzy logic are easily understood. Second, fuzzy logic has led to many successful applications, including many commercial products, in which the crucial part relies on representing and dealing with statements in natural language that involve vague terms. Third, fuzzy logic is a proper generalization of classical logic and, follows an agenda similar to that of classical logic, and has already been highly developed. An important consequence is that fuzzy logic extends the rich realm of applications of classical logic by applications in which the bivalent character of classical logic is a limiting factor.

Fuzzy Logic and Probability

Ever since the publication of (Zadeh 1965), the relationship between fuzzy logic and probability theory has been an object of another controversy. The various facets of this relationship have been discussed in many papers, including those contained in the special issues of *Computational Intelligence* (Vol. 4, No. 2, 1988), *IEEE Transactions on Fuzzy Systems* (Vol. 2, No. 1, 1994), and *Technometrics* (Vol. 37, No. 3, 1995). An extensive discussion on this topic comes as no surprise because both fuzzy logic and probability address the phenomenon of uncertainty and both use the real unit interval [0, 1]. The central questions of the debate include:

How does fuzzy logic relate to probability theory?
Is uncertainty the same as randomness?
Does the notion of probability exhaust all our notions of uncertainty?

The earliest paper discussing the relationship between fuzzy logic and probability is (Loginov 1966) in which the author suggests that membership degrees of fuzzy sets may be interpreted as conditional probabilities. This or a similar view has later been adopted by many people. Several leading researchers, including Cheeseman (1988a,b) and Lindley (1987), were repeatedly criticizing fuzzy logic on the ground that probability methods alone, and Bayesian methods in particular, are sufficient for representation and management of any type of uncertainty. As an illustration, the following is a quote from (Lindley 1987):

> The only satisfactory description of uncertainty is probability. By this I mean that every uncertainty statement must be in the form of a probability; that several uncertainties must be combined using the rules of probability; and that the calculus of probabilities is adequate to handle all situations involving uncertainty. ... We speak of “the inevitability of probability.”

In Sect. 16, Lindley concludes:

> ... probability is the only sensible description of uncertainty and is adequate for all problems involving uncertainty. All other methods are inadequate. ... My challenge that anything that can be done with fuzzy logic, ..., or any other alternative to probability, can better be done with probability, remains.

On the other hand, it has been pointed out many times, see e.g., (Klir 1989) and (Kosko 1990), that fuzzy logic studies a type of uncertainty that is fundamentally different from that studied by probability theory. As an example, take the proposition “Peter is a tall man.” As explained above, fuzzy logicians consider this as a many-valued (fuzzy) proposition, i.e., a proposition whose truth degree may be any degree from [0, 1] (or from another appropriate scale of truth degrees). The higher the degree, the truer the proposition. The graded nature of such propositions reflects the graded nature of human concepts such as the concept of a tall man. Note that the graded nature of human concepts was confirmed by many experiments in the psychology of concepts (Belohlavek and Klir 2011). Considering the proposition “Peter is a tall man.” as a bivalent proposition (yes-or-no proposition) is inadequate. For example, the question “Is the proposition true, but answer ‘yes’ or ‘no’ only?”, is inappropriate because it distorts the meaning of the concept of a tall man, namely it distorts its fuzziness. When probability theorists suggest that truth degrees of propositions are (conditional) probabilities, they assume that the propositions themselves are bivalent and that the truth degree measures a person’s (subjective) uncertainty of whether the proposition is true, i.e., whether the truth degree of the proposition is 1. Clearly, this view is very different from the view of fuzzy logicians. Because fuzzy propositions are considered bivalent in this view, the view is considered fundamentally inadequate by fuzzy logicians.

The above considerations point to the fact that fuzzy logic and probability study different types of uncertainty, that these types are complementary and are both important in human action. Hence, fuzzy logic and probability theory should be looked at as complementary rather than competitive theories. This situation was recognized in an early paper by Zadeh (1968). In order to extend the applicability of probability theory to account for fuzzy events such as “high inflation rate”, Zadeh proposed to generalize the concept of a probability space by allowing events to be fuzzy sets rather than ordinary sets of elementary events. The need for extensions of probability theory that take into account fuzziness of natural language expressions, which is particularly emphasized by the demand for natural language interfaces in web search, has recently been pointed out in several papers by Zadeh (2002, 2006).

In (Zadeh 2002) the following examples of simple problems are presented for which probability theory does not provide solutions:

Most Swedes are tall. Most Swedes are blond. What is the probability that a Swede picked at random is tall and blond?

Usually Robert returns from work at about 6 p.m. What is the probability that he is home at 6:30 p.m.?

A box contains about 20 balls of various sizes. A few are small and several are large. What is the probability that a ball drawn at random is neither large nor small?

In view of these examples, it becomes apparent that to base probability theory on bivalent logic results in a fundamental limitation and that, naturally, probability theory should be based on fuzzy logic. Such a conclusion presents a serious challenge for research in the foundations of probability theory.

Various Meanings of "Fuzzy Logic"

The term "fuzzy logic", coined by Goguen (1968), is used in several meanings. In its common-sense meaning, the term refers to formal and informal principles and methods of reasoning that involve vaguely defined concepts (concepts without clear-cut boundaries) that are based on the graded approach.

Two other meanings are frequently used, fuzzy logic in the narrow sense and fuzzy logic in the broad sense. Fuzzy logic in the narrow sense, called also mathematical fuzzy logic (Hájek 2006), develops deductive systems of logic very much in the style of classical mathematical logic. When the term fuzzy logic is used in the broad sense, it refers to an attempt to emulate human reasoning in natural language and includes aspects that are beyond the usual scope of mathematical logic. Fuzzy logic in the narrow and broader sense are discussed in more detail in section "Fuzzy Logic as Logic".

Basic Concepts of Fuzzy Logic

Truth Degrees and Truth Functions of Logical Connectives

As mentioned above, fuzzy logic uses a scale, denoted here by L, of truth degrees. A common choice for L is $[0, 1]$ (real unit interval) and unless stated otherwise, we assume $L = [0, 1]$ throughout this section. In general, L is usually assumed to be a complete lattice bounded by 0 and 1. As in classical logic (where $L = \{0, 1\}$), the scale needs to be equipped with (truth functions of) logical connectives such as conjunction, implication, etc. Unlike classical logic, where there truth functions are simply derived from the use of connectives in language and are unique (form example, "φ and ψ" is true if and only if both φ and ψ are true), fuzzy logic does not have unique truth functions of logical connectives. Namely, if there is no obvious way to define the truth degree of proposition "φ and ψ" given that the truth degree of φ and ψ are 0.7 and 0.8, respectively. Therefore, rather than defining a particular truth function of conjunction ("the right function"), fuzzy logic accepts as appropriate any truth function which satisfies certain conditions that come from intuitive requirements as well as from particular application contexts. For example, a truth function $\otimes$ of conjunction is a binary function $\otimes : L \times L \to L$ which needs to satisfy at least the following conditions:

$$a_1 \leq a_2 \text{ and } b_1 \leq b_2 \text{ implies } a_1 \otimes b_1 \leq a_2 \otimes b_2, \qquad \text{(monotonicity)}$$
$$a \otimes b = b \otimes a, \qquad \text{(commutativity)}$$
$$a \otimes (b \otimes c) = (a \otimes b) \otimes c, \qquad \text{(associativity)}$$
$$a \otimes 1 = 1 \otimes a = a, a \otimes 0 = 0 \otimes a = 0, \qquad \text{(boundary conditions)}$$

which are certainly intuitively appealing properties of conjunction. A function $\otimes$ on $L = [0, 1]$ satisfying these conditions is called a *t*-norm (Klement et al. 2000). The *t*-norms used in fuzzy logic are usually continuous (or at least left-continuous). The basic continuous *t*-norms are Gödel (maximum), Goguen (product), and Łukasiewicz *t*-norm, which are defined as follows:

$$\text{Gödel: } a \otimes b = \min(a, b), \tag{16.1}$$

$$\text{Goguen: } a \otimes b = a \cdot b, \tag{16.2}$$

$$\text{Łukasiewicz: } a \otimes b = \max(a + b - 1, 0). \tag{16.3}$$

Namely, any continuous *t*-norm can be obtained from the basic ones by so-called ordinal sum (Hájek 1998; Klement et al. 2000). *t*-norms have been extensively studied in the literature and various classes of *t*-norms, including classes of parameterized *t*-norms such as $a \otimes_\lambda b = 1 - \min\{1, [(1-a)_\lambda + (1-b)_\lambda]_\lambda\}$ for $\lambda \in [0, \infty)$ are described, e.g., in Gottwald (2001), Klement et al. (2000) and Klir and Yuan (1995).

In general, a truth function of an *n*-ary logical connective is a function $c : L^n \to L$. As in classical logic, further connectives such as disjunction, implication, or negation, are used in fuzzy logic. Due to limited scope we do not discuss the truth functions of these connectives here and refer the reader e.g., to Gottwald (2001) and Klir and Yuan (1995). An important question of a relationship between the truth functions of logical functions, such as the relationship between conjunction and implication, is discussed in section "Fuzzy Logic as Logic".

In addition to the connectives mentioned so far, fuzzy logic used various other connectives. For illustration, we mention linguistic modifiers and averaging functions. Modifiers are unary functions $m : [0,1] \to [0,1]$ which are thought of as the truth functions of unary connectives, called linguistic hedges Zadeh (1973, 1975), such as "very", "highly", "more or less", "somewhat", etc. Linguistic hedges are employed in linguistic rules such as "If temperature is very high, then …". A simple class of modifiers is given by

$$m_\lambda(a) = a_\lambda$$

for $a \in [0,1]$. For $a \in (0,1)$, the modifier is an increasing function and corresponds to linguistic hedges such as "more or less" or "somewhat". For $a \in (1,\infty)$, the modifier is a decreasing function and corresponds to intensifying linguistic hedges such as "very" or "highly". Averaging functions are defined as n-ary functions $c : [0,1]^n \to [0,1]$ that are non-decreasing, idempotent, and usually continuous and symmetric. Because they satisfy

$$\min(a_1,\ldots,a_n) \leq c(a_1,\ldots,a_n) \leq \max(a_1,\ldots,a_n)$$

and because min and max are "the largest (truth function of) conjunction" and "the least (truth function of) disjunction", averaging functions are thought of as filling a gap between conjunctions and disjunctions. As simple example is the arithmetical average $c(a,b) = \frac{a+b}{2}$. According to common sense, a person's financial wealth depends on whether his assets have good liquidity and his investments are good. Naturally, the degree $W(x)$ to which a person x is financially wealthy is obtained from the degrees $L(x)$ (good liquidity) and $I(x)$ (good investment) by means of an averaging function (e.g., $W(x) = \frac{L(x)+I(x)}{2}$) rather than a conjunction (e.g., $W(x) = \min\{L(x), I(x)\}$) or disjunction (e.g., $W(x) = \max\{L(x), I(x)\}$). Note that neither the modifiers nor the averaging functions have a counterpart in classical logic (modifiers are degenerate in classical logic, the only one is the identity function mapping 0 to 0 and 1 to 1; classical truth degrees cannot be averaged).

Fuzzy Sets and Fuzzy Relations

The concept of a fuzzy set generalizes the concept of a (characteristic function of a) classical set. A fuzzy set A in a universe U is defined as a mapping $A : U \to L$, i.e., A assigns to every element u from U a degree $A(u)$ from a scale L of truth degrees, called the degree of membership of u to A. If $L = [0,1]$, one usually speaks of standard fuzzy sets. Clearly, if $L = \{0,1\}$, we get the notion of a characteristic notion of an ordinary set.

The notions and operations related to fuzzy sets include both the counterparts of those from classical sets as well as new ones. An important example of the latter is the concept of an α-cut, which is defined for $\alpha \in L$ and a fuzzy set A as the ordinary subset ${}^{\alpha}A$ of U defined by ${}^{\alpha}A = \{u \in U \mid A(u) \geq \alpha\}$. A fuzzy set A is uniquely

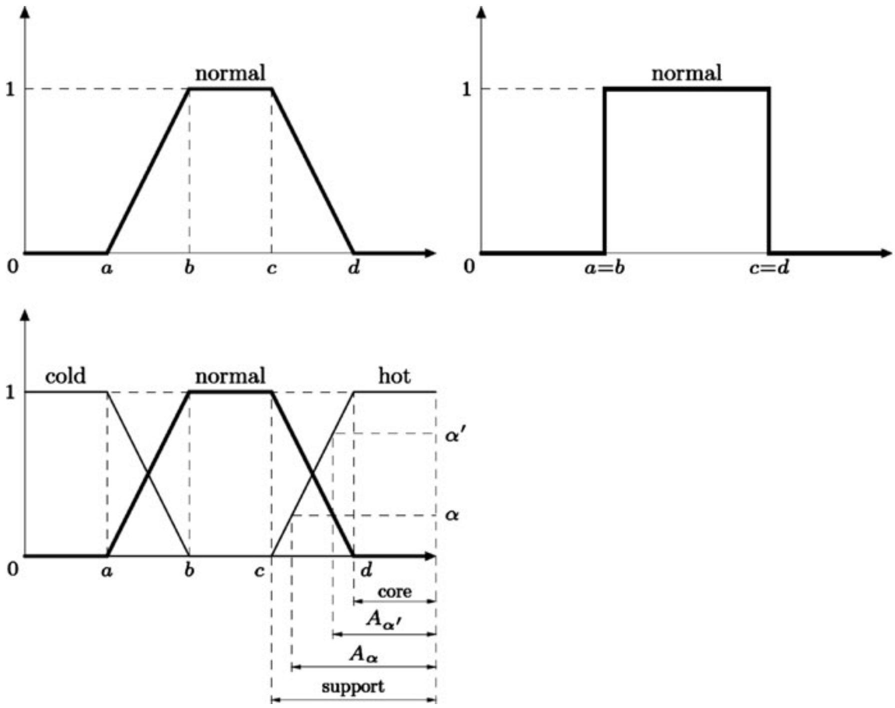


Fig. 16.1 Concept of fuzzy set

represented by the collection $\{{}^{\alpha}A \mid \alpha \in L\}$ of all of its α-cuts and this representation connects fuzzy sets with ordinary sets. The top part of Fig. 16.1 shows a fuzzy set representing the concept "normal" (temperature) versus a classical set representing the same concept. The bottom part shows three fuzzy sets, representing "cold", "normal", and "hot", and illustrates the concepts of an α-cut and support of a fuzzy set defined as $\mathrm{supp}(A) = \{u \in U \mid A(u) > 0\}$.

Every logical n-ary connective c on L induces a corresponding n-ary operation, defined component-wise. For example, if c is the truth function min of Gödel conjunction, the corresponding operation, called the standard intersection of fuzzy sets and denoted by $\cap$, is defined by

$$(A \cap B)(x) = \min(A(x), B(x)).$$

Relations on fuzzy sets can be both ordinary relations, such as the inclusion $\subseteq$ of fuzzy sets defined by $A \subseteq B$ if and only if $A(u) \leq B(u)$ for each $u \in U$. However, one may in general consider fuzzy versions of these relations, such as a degree of inclusion of fuzzy sets, which play an important role in fuzzy set theory.

Fuzzy relations are defined as fuzzy sets in Cartesian products. For example, a binary relation between sets U and V is a mapping $R : U \times V \to L$ with $R(u, v)$ being interpreted as a degree to which u is related to v. Among the several types of fuzzy relations used in applications, fuzzy equivalences (called also similarity relations) are perhaps the most important. A fuzzy relation $E : U \times U \to L$ is called a fuzzy equivalence if the following conditions generalizing the ordinary reflexivity, symmetry, and transitivity hold true:

$$E(u,u) = 1,$$
$$E(u,v) = E(v,u),$$
$$E(u,v) \otimes E(v,w) \leq E(u,w),$$

where $\otimes$ is a truth function of conjunction.

Various particular types of fuzzy sets and fuzzy relations are used in applications of fuzzy logic and were studied in the literature. Due to lack of space we omit details and refer the reader to numerous books on fuzzy sets and their applications, e.g., to Belohlavek (2002), Gottwald (2001), Klir and Yuan (1995) and Kruse et al. (1994).

Fuzzy Logic as Logic

Is there any logic in “fuzzy logic”, i.e., is it possible to develop a deductive system for reasoning which involves degrees of truth? What are the corresponding concepts of consequence, provability, completeness and what properties do they have? As was mentioned in section “What Is Fuzzy Logic?”, these question are addressed by fuzzy logic in the narrow sense. This section provides an introduction to the basic concepts involved.

Fuzzy Logic as Many-Valued Logic

Logics with more than two truth values, so-called many-valued logics, were studied in the field of mathematical logic since 1930s, see e.g., Gottwald (2001). Fuzzy logic can be considered a particular many-valued logic whose agenda is driven by the interpretation of truth values as truth degrees. Fuzzy logic uses many-valued counterparts of logical connectives of classical logic, as was discussed in section “Truth Degrees and Truth Functions of Logical Connectives”. In addition, fuzzy logic is truth functional. That is, if $\|\varphi\|$ and $\|\psi\|$ denote the truth degrees of formulas φ and ψ, the truth degree $\|\varphi \& \psi\|$ of the conjunction of φ and ψ is determined by

$$\|\varphi \& \psi\| = \|\varphi\| \otimes \|\psi\| \tag{16.4}$$

where $\otimes$ is a truth function of conjunction; and the same for other connectives.

Since in fuzzy logic, there are many possible choices of the truth functions of logical connectives (section "Truth Degrees and Truth Functions of Logical Connectives"), it is important to ask which combinations of truth functions are appropriate. An important argument regarding the choice of the truth functions of conjunction and implication comes from Goguen (1968) who showed that this question is connected to the rule of *modus ponens*. In particular, if one wants to have a good rule of *modus ponens* (yielding as much as possible but still sound), the truth functions $\otimes$ of conjunction and $\rightarrow$ of implication need to satisfy

$$a \otimes b \leq c \text{ if and only if } a \leq b \rightarrow c, \tag{16.5}$$

called the adjointness condition. For example, if $\otimes$ is a continuous (or even a left-continuous) t-norm, the unique $\rightarrow$ satisfying (16.5), called the residuum of $\otimes$, is given by

$$a \rightarrow b = \sup\{z \mid a \otimes z \leq b\}.$$

In particular, the residua of Gödel, Goguen, and Łukasiewicz t-norms, see (16.1)–(16.3), are given by

$$\text{Gödel}: a \rightarrow b = \begin{cases} 1 & \text{if } a \leq b, \\ b & \text{otherwise,} \end{cases}$$

$$\text{Goguen}: \mathrm{a} \rightarrow \mathrm{b} = \begin{cases} 1 & \text{if } a \leq b, \\ \frac{b}{a} & \text{otherwise,} \end{cases}$$

$$\text{Łukasiewicz}: a \rightarrow b = \min(1 - a + b, 1).$$

Ordinary-Style Calculi

Two basic types of fuzzy logical calculi can be distinguished. The first one are called ordinary-style calculi. Except for the fact that they allow more than two truth degrees, their its agenda is practically the same as that of classical logic. For example, formulas are defined as usual (starting from atomic formulas and applying logical connectives), a theory is a set of formulas, a proof from a theory T is a sequence of formulas which are either from T or result by application of a deduction rule to preceding formulas, etc. Due to truth functionality, the truth degree of a formula is defined as usual, cf. (16.4), given that particular structure $\mathbf{L}$ of truth degrees is chosen, i.e., a set L of truth degrees and truth functions of logical connectives from the language of the particular logical calculus. A tautology w.r.t. a class $\mathcal{L}$ of structures of truth degrees if for every structure $\mathbf{L} \in \mathcal{L}$, φ has truth degree 1 for every evaluation using truth degrees and logical connectives from $\mathbf{L}$. To illustrate ordinary-style completeness, consider the completeness theorem of propositional BL-logic Hájek (1998) that was proved in Cignoli et al. (2000). Given the axioms of BL-logic, the following conditions are equivalent for any formula φ:

1. φ is provable.
2. φ is a tautology w.r.t. the class of algebras which consist of [0, 1], a continuous t-norm, and its residuum.
3. φ is a tautology w.r.t. the class of BL-algebras (particular lattices equipped with operations $\otimes$ and $\rightarrow$, the algebraic counterparts of BL-logic).

For more information we refer to Gottwald (2001) and Hájek (1998).

Graded-Style (Pavelka-Style) Calculi

Graded-style calculi were introduced in a seminal paper by Pavelka (1979). Unlike ordinary-style calculi, the graded-style calculi works with formulas to which truth degrees are "attached". A pair $\langle \varphi, a \rangle$ carries a syntactical information that formula φ be true to degree at least a. For example, a theory is a set consisting of such pairs $\langle \varphi, a \rangle$ which specify that φ is assumed to be true to degree at least a. A deductive rule has two components, one working on formulas, the other working on truth degrees. For example, the rule of *modus ponens* applied to $\langle \varphi \Rightarrow \psi, a \rangle$ and $\langle \varphi, b \rangle$ yields a pair $\langle \psi, a \otimes b \rangle$ and reads as follows: If $\varphi \Rightarrow \psi$ and φ are true to degree at least a and b, respectively, ψ is true to degree at least $a \otimes b$. One then introduces the concept of a degree $|\varphi|_T$ to which formula φ is provable from theory T (supremum of as over all $\langle \varphi, a \rangle$ which can be obtained from the axioms and T using deduction rules) and the concept of a degree $\|\varphi\|_T$ to which φ is (semantically) entailed by T (infimum of truth degrees of φ in all models of T). A completeness theorem then says

$$|\varphi|_T = \|\varphi\|_T,$$

i.e., degree of probability equals degree of entailment. For further information including various particular graded-style calculi we refer to Belohlavek and Vychodil (2005, 2006), Gerla (2001) and Hájek (1998).

Fuzzy Logic in a Broad Sense

Note that from a general viewpoint of logic as a discipline studying human reasoning, fuzzy logic in the broad sense also fits the picture of fuzzy logic as logic. As mentioned in section "What Is Fuzzy Logic?", fuzzy logic in the broad sense attempts to emulate human reasoning. Conceptually, fuzzy logic in the broad sense is being developed in numerous papers by Zadeh (1973, 1975, 1979, 2006, 2008). Parts of fuzzy logic in the broad sense are highly developed and have numerous applications, for example the rule-based systems employed in fuzzy control, discussed in sections "Fuzzy Logic and Control" and "Success of Mamdani Control in Automobile Industry".

Note however, that traditional logical aspects of logic are as a rule of little concern in those developments, but see Hájek's chapter on logical analysis of the compositional rule of inference in (Hájek 1998) and also (Novák et al. 1999). From this point of view, fuzzy logic in the broad sense is at an early stage of development.

Fuzzy Logic and Control

The biggest success of fuzzy logic in the field of industrial and commercial applications has been achieved with *fuzzy controllers*. Fuzzy control is a way of defining a nonlinear table-based controller whereas its nonlinear transition function can be defined without specifying every single entry of the table individually. Fuzzy control does not result from classical control engineering approaches. In fact, its roots can be found in the area of rule-based systems. Fuzzy controllers simply comprise a set of vague rules that can be used for knowledge-based interpolation of a vaguely defined function.

Suppose we consider a technical system. For this system, we dictate a desired behavior. Generally a time-dependent *output variable* must reach a desired set value. The output is influenced by a *control variable* which we can manipulate. Finally, there exists a time-dependent *disturbance variable* that influences the output as well. The current control value is usually determined based on the current measurement values of the output variable ξ, the variation of the output $\Delta\xi = \frac{d\xi}{dt}$ and further variables.

Hereafter we will refer to input variables $\xi_1 \in X_1, \ldots, \xi_n \in X_n$ and one control variable $\eta \in Y$. The solution of a control problem is a suitable control function $\varphi : X_1 \times \ldots \times X_n \rightarrow Y$ that determines an appropriate control value $y = \varphi(\mathbf{x})$ for every input tuple $\mathbf{x} = (x^{(1)}, x^{(2)}, \ldots, x^{(n)}) \in X_1 \times \ldots \times X_n$. In classical control engineering, φ is commonly determined by solving a set of differential equations. It is very often out of the question to specify an exact set of differential equations. Note that human beings, however, are greatly able to control certain processes without knowing about higher mathematics.

Simulating the behavior of a human "controller" can be done by questioning the individual directly. An alternative would be extract essential information by observing the controlled process. The result of such *knowledge-based analysis* is a set of *linguistic rules* that control the process. Linguistic rules comprise a premise and a conclusion. The former relates to a fuzzy description of the crisp measured input, where the latter defines a suitable fuzzy output. Thus we need to formalize mathematical descriptions of the linguistic expressions used in the rules. Furthermore initialized rules need to be accumulated to result in one fuzzy output value. Finally, a crisp output value must be computed from the fuzzy one. The whole architecture for that knowledge-based model of a fuzzy controller is shown in Fig. 16.2.

The *fuzzification interface* operates on the current input value $\mathbf{x}_0$. If needed, $\mathbf{x}_0$ is mapped into a suitable domain, e.g., normalization to the unit interval. It also

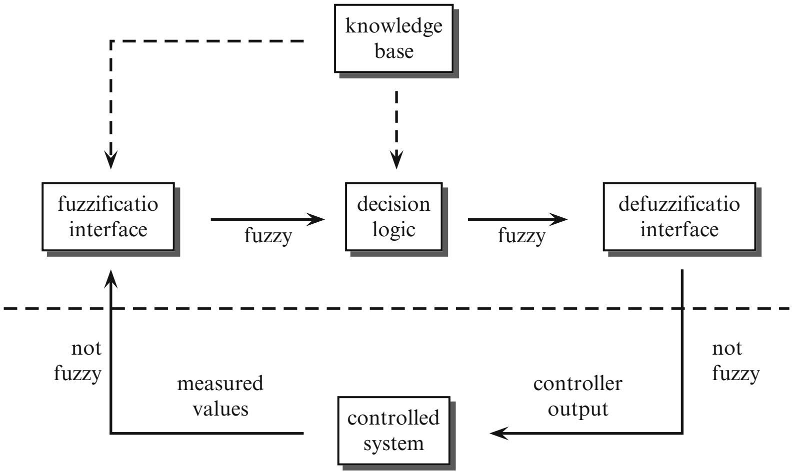


Fig. 16.2 Architecture of a fuzzy controller

transforms $\mathbf{x}_0$ into a linguistic term or fuzzy set. The *knowledge base* comprises the *data base*, i.e., all pieces of information about variable ranges, domain transformations, and the fuzzy sets with their corresponding linguistic terms. Moreover, it also contains a *rule base* storing the linguistic rules for controlling. The *decision logic* determines the output value of the corresponding measured input using the knowledge base. The *defuzzification interface* produces the crisp output value given the fuzzy output.

There exist two fundamentally different approaches to fuzzy control. Both of them are motivated intuitively (see the next two sections). We will see in section "Approximate Reasoning" that a fuzzy controller based on logical implications results in completely different methods of computation.

Mamdani-Assilian Controller

In 1975, the first model of a fuzzy controller was created by Ebrahim "Abe" Mamdani and his student Sedrak Assilian (Mamdani and Assilian 1975). Mamdani and Assilian developed their idea application-driven to control a steam engine based on human expert knowledge.

Here, the knowledge of an expert must be expressed by linguistic rules. First, for the set X_1, p_1 fuzzy sets $\mu_1^{(1)}, \ldots, \mu_{p_1}^{(1)} \in \mathcal{F}(X_1)$ must be defined. Accordingly, each fuzzy set is named with a suitable linguistic term. Second, X_1 is partitioned by its fuzzy sets. To be able to interpret each fuzzy set as fuzzy value or fuzzy interval, it is favorable to only use unimodal membership functions. Also, fuzzy sets of one partition should be disjoint, i.e., they satisfy

$$i \neq j \Rightarrow \sup_{x \in X_1} \left\{ \min\left\{ \mu_i^{(1)}(x), \mu_j^{(1)}(x) \right\} \right\} \leq 0.5.$$

Having divided X_1 into p_1 fuzzy sets $\mu_1^{(1)}, \ldots, \mu_{p_1}^{(1)}$, we partition the remaining sets $X_2, \ldots, X_n$ and Y in the same manner. Finally, these fuzzy partitions and the linguistic terms associated with the fuzzy sets correspond to the data base in our knowledge base.

The rule base is specified by rules of the form

$$\textbf{if } \xi_1 \textbf{ is } A^{(1)} \textbf{ and} \ldots \textbf{and } \xi_n \textbf{ is } A^{(n)} \textbf{ then } \eta \textbf{ is } B \tag{16.6}$$

whereas $A^{(1)}, \ldots, A^{(n)}$ and B represent linguistic terms corresponding to fuzzy sets $\mu^{(1)}, \ldots, \mu^{(n)}$ and μ, respectively, according to fuzzy partitions of $X_1 \times \ldots \times X_n$ and Y. Hence the rule base comprises k control rules

$$R_r : \textbf{if } \xi_1 \textbf{ is } A_{i_{1,r}}^{(1)} \textbf{ and} \ldots \textbf{and } \xi_n \textbf{ is } A_{i_{n,r}}^{(n)} \textbf{ then } \eta \textbf{ is } B_{i_r}, \quad r = 1, \ldots, k.$$

Remark that these rules are not regarded as logical implications. They rather define $\eta = \varphi(\xi_1, \ldots, \xi_n)$ piecewise where

$$\eta \approx \begin{cases} B_{i_1} & \text{if } \xi_1 \approx A_{i_{1,1}}^{(1)} \text{ and } \ldots \text{ and } \xi_n \approx A_{i_{n,1}}^{(n)}, \\ \vdots & \vdots \\ B_{i_k} & \text{if } \xi_1 \approx A_{i_{1,k}}^{(1)} \text{ and } \ldots \text{ and } \xi_n \approx A_{i_{n,k}}^{(n)}. \end{cases}$$

Since the rules are treated as *disjunctive*, we can say that the control function φ is obtained by knowledge-based interpolation.

Observing a measurement $\mathbf{x} \in X_1 \times \ldots \times X_n$ the decision logic applies each R_r separately. It computes the degree to which $\mathbf{x}$ fulfills the premise of R_r, i.e., the degree of applicability

$$\alpha_r \stackrel{\text{def}}{=} \min\left\{ \mu_{i_{1,r}}^{(1)}(x^{(1)}), \ldots, \mu_{i_{n,r}}^{(n)}(x^{(n)}) \right\}. \tag{16.7}$$

"Cutting off" the output fuzzy set μ_{i_r} of rule R_r at α_r leads to the rule's output fuzzy set:

$$\mu_{\mathbf{x}}^{\mathrm{o}(R_r)}(y) = \min\left\{ \alpha_r, \mu_{i_r}(y) \right\}. \tag{16.8}$$

Having computed all α_r for $r = 1, \ldots, k$, the decision logic combines all $\mu_{\mathbf{x}}^{\mathrm{o}(R_r)}$ applying the t-conorm maximum in order to get the overall output fuzzy set

$$\mu_{\mathbf{x}}^{\mathrm{o}}(y) = \max_{r=1,\ldots,k} \left\{ \min\left\{ \alpha_r, \mu_{i_r}(y) \right\} \right\}. \tag{16.9}$$

In control engineering, a crisp control value is needed. Therefore $\mu_{\mathbf{x}}^{\mathrm{o}}$ is forwarded to the defuzzification interface. Here, it depends on the kind of method that is implemented to defuzzify $\mu_{\mathbf{x}}^{o}$. The most well-known approaches are the max criterion method, the mean of maxima (MOM) method and the center of gravity (COG) method. Using the first approach, simply an arbitrary value $y \in Y$ is chosen for which $\mu_{\mathbf{x}}^{o}(y)$ reaches a maximum membership degree. Picking a random value leads to a nondeterministic control behavior which is usually undesired. The MOM method choses the mean value of the set of elements $y \in Y$ resulting in maximal membership degrees. The defuzzified control value η might not even be in the set which can lead to unexpected control actions. The COG method defines the value located under the center of gravity of the area $\mu_{\mathbf{x}}^{o}$ as control value η, i.e.,

$$\eta = \left(\int_{y \in Y} \mu_{\mathbf{x}}^{o}(y) \cdot y \, \mathrm{d}y\right) \Big/ \left(\int_{y \in Y} \mu_{\mathbf{x}}^{o}(y) \, \mathrm{d}y\right). \tag{16.10}$$

In most control applications, this method shows smooth control behaviors. However, it might even lead to counterintuitive results as well. For a more profound discussion about defuzzification, see e.g., Kruse et al. (1994).

Let us conclude this type of controller by analyzing the form of linguistic rules again. Regarding (16.8), it is clear that the minimum is used as fuzzy implication. Obviously this does not coincide with its crisp counterpart. Just consider $p \rightarrow q$ knowing that p is false. Then $p \rightarrow q$ is true regardless of the truth value of q in classical propositional logic. However, $\min\{0, q\}$ is always 0. One way to justify the heuristic of Mamdani and Assilian is to replace the concept of implication by the one of *association* (Cordón et al. 1999). We say that for a rule R_r an output fuzzy set B_{i_r} is associated with n input fuzzy sets $A_{i_{j,r}}^{(j)}$ for $j = 1, \ldots, n$. This association is modeled by a fuzzy conjunction, e.g., the t-norm min.

We retrieve Mamdani's heuristics by extensionality assumptions (Klawonn et al. 1995; Klawonn and Kruse 1993). If the fuzzy relation R relating the $x^{(j)}$ and y satisfies some extensionality properties, then Mamdani's approach is derived in the same way. Let E and E' be two similarity relations defined on the domains X and Y of x and y, respectively. The extensionality of R on $X \times Y$ thus means

$$\begin{aligned} &\forall x \in X : \forall y, y' \in Y : R(x, y) \otimes E'(y, y') \leq R(x, y'), \\ &\forall x, x' \in X : \forall y \in Y : R(x, y) \otimes E(x, x') \leq R(x', y). \end{aligned} \tag{16.11}$$

So, if $(x, y) \in R$, then x will be related to the neighborhood y. The same shall hold for y in relation to x. Then $A_r^{(j)}(x) = E(x, a_r^{(j)})$ and $B_r(x) = E'(y, b_r)$ can be seen as fuzzy sets of values that are close to $a_r^{(j)}$ and b_r, respectively. Naturally, $\forall r = 1, \ldots, k : R\left(a_r^{(1)}, \ldots, a_r^{(p)}), b_r\right) = 1$. The user thus only needs to define reasonable similarity relations E_j and E' for each input ξ_j and the output η, respectively. Then, using the extensionality properties of R, one gets

$$R(x^{(1)}, \ldots, x^{(p)}, y) \geq \max_{r=1, \ldots, k} \otimes \left(A_r^{(1)}(x^{(1)}), \ldots, A_r^{(p)}(x^{(p)}), A_r(y)\right).$$

If we use the t-norm $\otimes = \min$, then Mamdani's approach to compute the fuzzy output is obtained. In (Boixader and Jacas 1998; Klawonn and Castro 1995) indistinguishability or similarity is expressed as link between the extensionality property and fuzzy equivalence relations. Fuzzy interpolation can be also seen as logical inference given fuzzy information coming from an vaguely known function (Klawonn and Novák 1996). Likewise, in Sudkamp (1993) fuzzy rules are obtained from set of pairs (a_i, b_i) and similarity relations on X and Y.

Takagi-Sugeno Controller

Takagi-Sugeno controllers (Takagi and Sugeno 1985) can be seen as modification of Mamdani-Assilian controllers. For both controllers, we need to specify fuzzy partitions of the input domains. However, no fuzzy partition of the output domain is needed since the rules R_r for $r = 1, \ldots, k$ are given as

$$R_r : \textbf{if } \xi_1 \textbf{is } A^{(1)}_{i_{1,r}} \textbf{ and} \ldots \textbf{and } \xi_n \textbf{ is } A^{(n)}_{i_{n,r}} \textbf{ then } \eta = f_r(\xi_1, \ldots, \xi_n).$$

Usually the functions f_r are linear, i.e., $f_r(\mathbf{x}) = a_r^{(0)} + \sum_{i=1}^{n} a_r^{(i)} x^{(i)}$.

Again, the decision logic determines the degree of applicability α_r of each premise using (16.7). These degrees are directly used to determine a crisp control value

$$\eta = \frac{\sum_{r=1}^{k} \alpha_r \cdot f_r(\mathbf{x})}{\sum_{r=1}^{k} \alpha_r}$$

which is a weighted sum over all rules' outputs. Hence, the defuzzification is omitted for that type of controller.

Approximate Reasoning

So far, we have treated the linguistic rules as associations of an n-dimensional fuzzy input point with one fuzzy output. This makes sense for control applications where each rule defines an operating point of the system to be controlled. Another way to interpret a fuzzy controller is to fuzzy constrain the control function by the fuzzy rules. This can be done by interpreting the inference process as approximate reasoning. In classical reasoning, tautologies/inference rules are used for deductive inferences of crisp conclusions from *crisp* propositions. Approximate reasoning can be seen as generalization of classical reasoning applied to *fuzzy* propositions. In (Zadeh 1973), first approaches have been developed to generalize approximate reasoning to fuzzy sets. In (Zadeh 1979, 1983), this methodology is explained in more detail. Using possibility distributions to represent incomplete knowledge helps to understand the mention techniques.

Whereas fuzzy set theory is closely associated with vague concepts, the application of possibility theory (Dubois and Prade 1988) relates to the imperfect description of an existing element x_0 in a set $A \subseteq X$. Possibility theory can be seen as counterpart to probability theory. In order to describe a possibility distribution $\Pi : 2^X \to [0,1]$, the following axioms are used:

$$\begin{aligned} &\Pi(\emptyset) = 0, \\ &\Pi(A) \leq \Pi(B) \text{ if } A \subseteq B \text{ and} \\ &\Pi(A \cup B) = \max\{\Pi(A), \Pi(B)\} \text{ for all } A, B \subset X. \end{aligned}$$

$\Pi(A) = 1$ means that $x_0 \in A$ is unconditional possible. If $\Pi(A) = 0$ then it is impossible that $x_0 \in A$. In Zadeh (1978), uncertainty about x_0 is modeled by the possibility measure $\Pi : 2^\Omega \to [0,1], \Pi(A) = \sup\{\mu(x) \mid x \in A\}$ when a fuzzy set $\mu : x \to [0,1]$ is given as only description of x_0. For this special case the possibility measure is given by the possibility degrees of the singletons, i.e., $\Pi(\{x\}) = \mu(x)$.

For simplicity consider one-dimensional input and output spaces, respectively. Here, the choice of an appropriate two-dimensional possibility distribution is crucial. The rule

$$R : \textbf{if } \xi \textbf{ is } A \textbf{ then } \eta \textbf{ is } B$$

that associates the input fuzzy set μ_A with the output fuzzy set μ_B is modeled by a possibility distribution

$$\pi_{X,Y}(x,y) = I(\mu_A(x), \mu_B(y))$$

whereas I is an implication of a multivalued logic. Hence $\mu_B = \mu_A \circ \pi_{X,Y}$ where $\pi_{X,Y}$ is a fuzzy relation on $X \times Y$. The composition of a fuzzy set μ with a fuzzy relation π is defined by

$$\mu \circ \pi : Y \to [0,1], \quad y \mapsto \sup_{x \in X} \{\min\{\mu(x), \pi(x,y)\}\}.$$

This is clearly a fuzzification of the composition $\circ$ of two crisp sets $M \subseteq X$ and $R \subseteq X \times Y$, i.e.,

$$M \circ R \stackrel{\text{def}}{=} \{y \in Y \mid \exists x \in X : (x \in M \wedge (x,y) \in R)\} \subseteq Y.$$

The task in fuzzy control based on such relational equations is to find a fuzzy relation π that fulfills all equations $\mu_{B_r} = \mu_{A_r} \circ \pi$ for every rule R_r with $r = 1, \ldots, k$. If multiple inputs $X_1, \ldots, X_n$ are used, then μ_A is defined on the product space $X = X_1$

$\times \ldots \times X_n$ as in (16.7). For each of the k relational equations, the *Gödel relation* is determined by

$$(x, y) \in \pi^G_{X,Y} \Longleftrightarrow (x \in \mu_A \rightarrow y \in \mu_B)$$

where the implication $\rightarrow$ is evaluated by the Gödel implication (see section "Fuzzy Logic as Many-Valued Logic"). Thus a linguistic rule can be seen as gradual rule 'The more μ_A, the more μ_B' which constrains π by the inequality

$$\min(\mu_A(x), \pi(x, y)) \leq \mu_B(y)$$

for all $(x, y) \in X \times Y$. Theoretically, different fuzzy implications could be used to describe π. However, several reasons can be found in favor for I_G, e.g., Dubois and Prade (1985, 1992).

If the system of relational equations $\mu_{B_r} = \mu_{A_r} \circ \pi$ for $r = 1, \ldots, k$ is solvable, then

$$\pi^G = \bigcap_{r=1}^{k} \pi^G_r(\mu_{A_r}(x), \mu_{B_r}(y))$$

is a solution with $\cap$ being the minimum t-norm. At the same time this is the greatest solution. We can say that the relation $\prod(\{(x, y)\}) \stackrel{\text{def}}{=} \pi(x, y)$ gives an estimate whether it is *possible* that input tuple x is assigned to output value y. So, the set of *conjunctive* rules imposes *soft constraints* on the control function φ. In practice, these constraints may lead to contradictions if narrow output fuzzy sets with overlapping input fuzzy sets are used. Thus the controller would output the empty fuzzy set, i.e., no solution. It is therefore reasonable to define rather narrow fuzzy sets for the input variables and rather broader fuzzy sets for the output.

Success of Mamdani Control in Automobile Industry

In the 1990s many real-world control applications have been greatly solved using Mamdani's approach. Among them are many control problems in the industrial automobile field. The number of publications, however, is really low. Two control applications at Volkswagen AG successfully use Mamdani's approach, i.e., the engine idle speed control and the shift-point determination of an automatic transmission (Schröder et al. 1997). The idle speed controller is based on similarity relations (see section "Mamdani-Assilian Controller"). This helps to view the control function as interpolation of a point-wise known function. The shift-point determination continuously adapts the gearshift schedule between two extremes, i.e., economic and sporting. A sport factor is computed to individually adapt the gearshift movements of a driver.

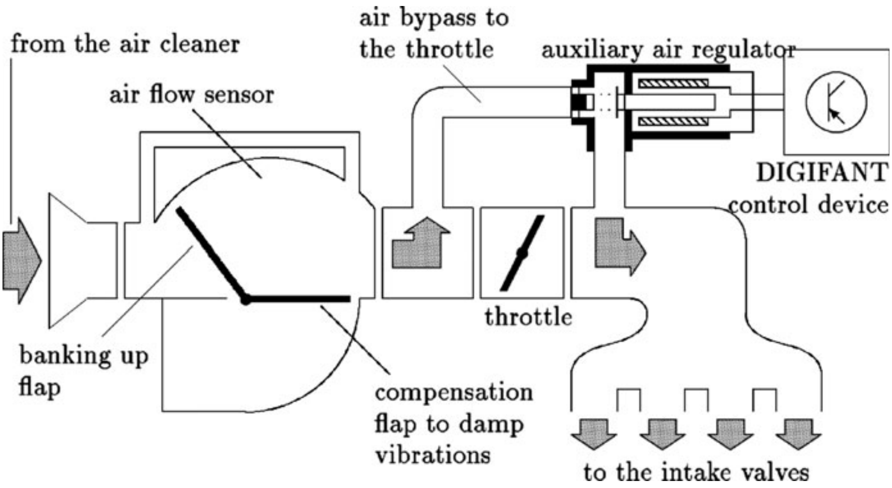


Fig. 16.3 Principle of the engine idle speed control

Engine Idle Speed Control

The task is to control the idle speed of a spark ignition engine. One way is a volumetric control where an auxiliary air regulator alters the cross-section of a bypass to the throttle. This is depicted in Fig. 16.3.

The pulse width of the auxiliary air regulator is changed by the controller. If there is a drop in the number of revolutions, then the controller forces the auxiliary air regulator to increase the bypass cross-section. The air flow sensor measures the increased air flow rate and thus notifies the controller. The new quantity for the fuel injection must be computed. Due to a higher air flow rate, the engine yields more torque. This again results in a higher number of revolutions which could be reduced analogously by decreasing the bypass cross-section.

Both fuel consumption and pollutant emissions should be ultimately reduced. This can be reached by slowing down the idle speed. However, a switching on of certain automobile facilities, e.g., air-conditioning system, forces the number of revolutions to drop. Hence the controller must be very flexible. More problems involved in this control application can be found in Schröder et al. (1997).

Due to this motivating problem, a Mamdani fuzzy controller was developed based on similarity relations. The resulting fuzzy controller was easier to design and showed an improved control behavior compared to classical control approaches. Similarity relations to represent indistinguishability or similarity of points within a certain vicinity seems to be a natural modeling way for engineers.

In fact, indistinguishability is not produced by measurement errors or deviations. It just expresses that arbitrary precision is not necessary to control a system. A control expert must thus specify a set of k input-output tuples $((x_r^{(1)}, \ldots, x_r^{(p)}), y_r)$. For each $r = 1, \ldots, k$, the output value y_r seems appropriate for the input $(x_r^{(1)}, \ldots, x_r^{(p)})$. So, the human expert defines the partial control function φ_0.

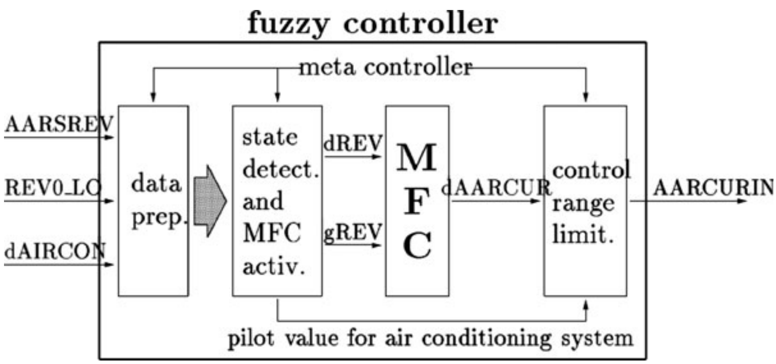


Fig. 16.4 Structure of the fuzzy controller

In the 1990s the question to be answered was to compute a suitable output value for an arbitrary input given specified similarity relations and φ_0 (Schröder et al. 1997). Using the extensionality properties defined in (16.11), one obtains Mamdani's fuzzy output directly by computing the extensional hull of φ_0 given the similarity relations. The partial control function φ_0 can thus be reinterpreted as k control rules of the form:

$$R_r:\ \textbf{if}\ \xi_1\ \textbf{is}\ \textit{approximately}\ x_r^{(1)}\ \textbf{and}\ \ldots\ \textbf{and}\ \xi_p\ \textbf{is}\ \textit{approximately}\ x_r^{(p)}\ \textbf{then}\ \eta\ \textbf{is}\ \textit{approximately}\ y_r.$$

A more profound theoretical analysis of this approach can be found in Klawonn et al. (1995).

To control the engine idle speed controller, two input variables are needed:

1. The deviation dREV [rpm] of the number of revolutions to the set value, and
2. The gradient gREV [rpm] of the number of revolutions between two ignitions.

The only output variable is the change of current dAARCUR for the auxiliary air regulator. The controller is shown in Fig. 16.4.

The knowledge to control the engine idle speed controller was extracted by measurement data obtained from idle speed experiments. The partial control mapping $\varphi_0 : X_{(\mathrm{dREV})} \times X_{(\mathrm{gREV})} \to Y_{(\mathrm{dAARCUR})}$ has been specified as in Table 16.1 (left-hand side).

Using a similarity relation and φ_0, the fuzzy controller was defined. Its induced control surface is shown in Fig. 16.5 as a grid of supporting points. The center of area (COA) method has been used for defuzzification. To obtain the corresponding Mamdani fuzzy controller, each point of φ_0 was associated with a linguistic term, e.g., negative big (nb), negative medium (nm), negative small (ns), approximately zero (az), and so on. The obtained fuzzy partitions of all three variables are shown in Figs. 16.6–16.8, respectively. The partial mapping φ_0 was translated into linguistic rules of the form

if dREV **is** A **and** gREV **is** B **then** dAARCUR **is** C.

Table 1.1 The partial control mapping ∏ 0 (left-hand side) and its corresponding fuzzy rule base (right-hand side).

		gREV						
		–40	–6	–3	0	3	6	40
	–70	20	15	15	10	10	5	5
	–50	20	15	10	10	5	5	0
	–30	15	10	5	5	5	0	0
dREV	0	5	5	0	0	0	–10	–5
	30	0	0	0	–5	–5	–10	–15
	50	0	–5	–5	–10	–15	–15	–20
	70	–5	–5	–10	–15	15	15	15

		gREV						
		nb	nm	ns	az	ps	pm	pb
	nb	ph	pb	pb	pm	pm	ps	ps
	nm	ph	pb	pm	pm	ps	ps	az
	ns	pb	pm	ps	ps	az	az	az
dREV	az	ps	ps	az	az	az	nm	ns
	ps	az	az	az	ns	ns	nm	nb
	pm	az	ns	ns	ns	nb	nb	nh
	pb	ns	ns	nm	nb	nb	nb	nh

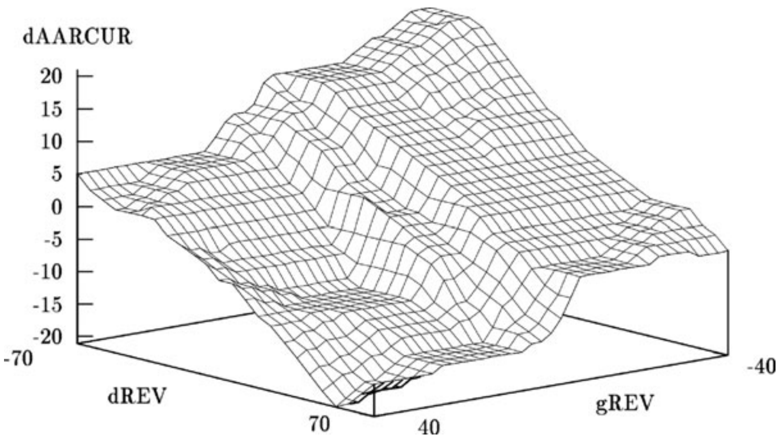
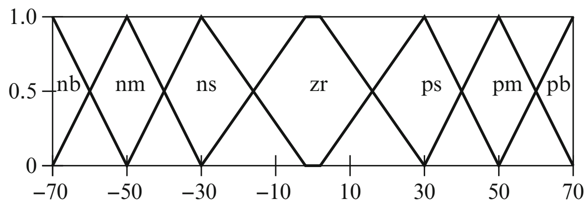


Fig. 16.5 Performance characteristics

Fig. 16.6 Deviation dREV of the number of revolutions



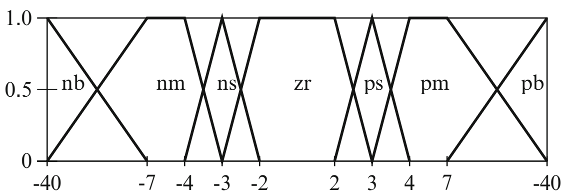


Fig. 16.7 Gradient gREV of the number of revolutions

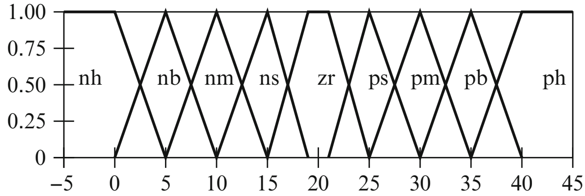


Fig. 16.8 Change of current dAARCUR for the auxiliary air regulator

The complete set of rules is given on the right-hand side of Table 16.1.

In (Klawonn et al. 1995; Schröder et al. 1997) the Mamdani fuzzy controller shows a very smooth control behavior compared to its serial counterpart. Furthermore the fuzzy controller reaches the desired set point precisely and fast. Its behavior is robust even with slowly increasing load. Thus the number of revolutions does not lead to any vibration even after extreme changes of load occur.

Flowing Shift-Point Determination

Conventional automatic transmissions select gears based on so-called gearshift diagrams. Here, the gearshift simply depends on the accelerator position and the velocity. A lagging between up and down shift avoids oscillating gearshift when the velocity varies slightly, e.g., during stop-and-go traffic. For a standardized behavior, a fixed diagram works well. Until 1994, the Volkswagen gear box had two different types of gearshift diagrams, i.e., economic "ECO" and sporting "SPORT". An economic gearshift diagram switches gears at a low number of revolutions to reduce the fuel consumption. A sporting one leads to gearshifts at a higher number of revolutions. Since 1991 it was a research issue at Volkswagen AG to develop an individual adaption of shift-points. No additional sensors should be used to observe the driver.

The idea was that the car "observes" the driver (Schröder et al. 1997) and classifies him or her into calm, normal, sportive (assigning a sport factor $\in [0, 1]$), or nervous (to calm down the driver). A test car from Volkswagen was operated by many different drivers. These people were classified by a human expert (passenger). Simultaneously, 14 attributes were continuously measured during test drives. Among them were variables like the velocity of the car, the position of the acceleration pedal, the speed of the acceleration pedal, the kick down, or the steering wheel angle.

The final Mamdani controller was based on four input variables and one output. The basic structure of the controller is shown in Fig. 16.9. In total, 7 rules could be

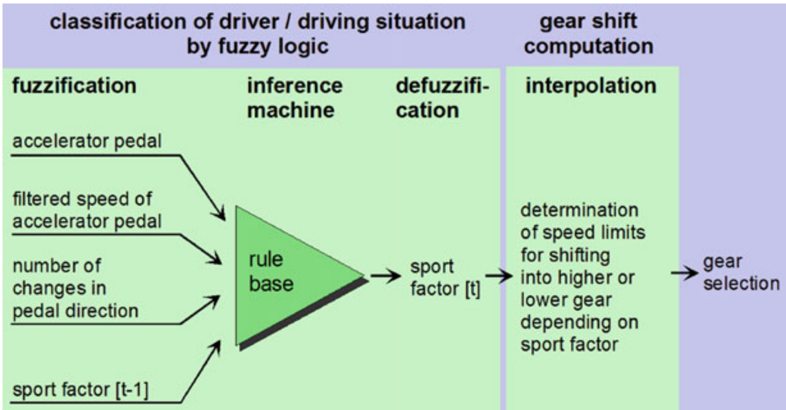


Fig. 16.9 Flowing shift-point determination with fuzzy logic

identified at which the antecedent consists of up to 4 clauses. The program was highly optimized: It used 24 Byte RAM and 702 Byte ROM, i.e., less than 1 KB. The runtime was 80 ms which means that 12 times per second a new sport factor was assigned. The controller is in series since January 1995. It shows an excellent performance.

Fuzzy Logic and Knowledge Discovery in Databases

Knowledge discovery in databases (KDD) tries to inspect, clean, transform and model *data* in large databases in order to find useful information or support decision making. Ultimately, one tries to formulate knowledge based on pieces of information that have been discovered in databases. A single datum may describe the condition of a certain object. It carries only information if there are at least two different states of the condition. A datum might be seen as the realization of a certain variable of a universe. There are different representations of a datum as it has been measured, i.e., nominal, ordinal, interval or ratio (Stevens 1946).

The KDD process is usually performed in four stages. At the first stage, the data are valuated and examined w.r.t. simple and essential characteristics, e.g., analysis of frequency, reliability test, runaway, credibility. The second stage comprises pattern matching or the grouping of observations. Usually transformations are performed with the goal to find structures within data. At that stage, exploratory data analysis is performed to examine the data without a previously chosen mathematical model. At the third level, data are analyzed w.r.t. one or more mathematical models. These models can be either qualitative or quantitative. The former one is the formation relating to additional characteristics expressed by quality, e.g., introduction of the term of similarity for cluster analysis. The latter type of models tries to recognize functional relations, e.g., an approximation of regression analysis.

At the fourth level, conclusions from the whole process are drawn and evaluated. Also, future or missing values might be predicted. Sources of data may be combined by, e.g., data fusion. In general, the learned models are revised at that stage.

If data are vague, imprecise or inconsistent, the application of fuzzy logic to KDD might improve results. Usually common data are analyzed by fuzzy methods whereas some researchers also analyze fuzzy data. The most prominent approach to fuzzy data analysis is fuzzy clustering that is introduced in section "Fuzzy Clustering". Its successfulness in KDD might come from the fact that human beings do not group objects based on crisp labels. We rather use some kind of fuzzy terms to cluster things, e.g., into the group of tall people. Many everyday decisions are fuzzy and human beings are able to handle that. Therefore an appropriate answer to the following question is naturally important: How can a computer learn fuzzy rules from data to explain or support decisions like people do? We describe some general approaches to generate fuzzy rules from data in section "Fuzzy Rule Generation".

Fuzzy Clustering

Clustering is an unsupervised learning task that tries to divide data s.t.

- Objects belonging to the same cluster are as similar as possible, and
- Objects belonging to different clusters are as dissimilar as possible.

Similarity is normally measured in terms of a distance function. The smaller the distance, the more similar two data tuples. Here, we assume that every data tuple is an element of the n-dimensional Euclidean space IR^n.

Definition 1 (Distance function). *The mapping* $d : \mathrm{IR}^n \times \mathrm{IR}^n \rightarrow [0,\infty)$ *is a distance function if it satisfies the following conditions for all* $\mathbf{x},\mathbf{y},\mathbf{z} \in \mathrm{IR}^n$:

1.	$d(\mathbf{x},\mathbf{y}) = 0 \Leftrightarrow \mathbf{x} = \mathbf{y}$	(identity),
2.	$d(\mathbf{x},\mathbf{y}) = d(\mathbf{y},\mathbf{x})$	(symmetry),
3.	$d(\mathbf{x},\mathbf{z}) \leq d(\mathbf{x},\mathbf{y}) + d(\mathbf{y},\mathbf{z})$	(triangle inequality).

Henceforth we only focus on partitioning algorithms, i.e., given a number $c \in \mathrm{IN}$, find the best partition of data into c groups. This is fundamentally different from hierarchical clustering techniques where data are organized in a nested sequence of groups (e.g., dendrograms). Usually the true number of clusters is unknown which makes it hard to use partitioning methods. To further specify, we concentrate on *prototype-based clustering* algorithms where clusters are represented by prototypes $C_i, i = 1, \ldots, c$. The prototypes shall capture the structure/distribution of data in each cluster. They are constructed by clustering algorithms. For simplicity, consider cluster prototypes C_i which are solely represented by the cluster centers $\mathbf{c}_i$. Furthermore, the distance measure d is based on the inner product, e.g., the Euclidean distance

$$d(\mathbf{x},\mathbf{y}) = \|\mathbf{x}-\mathbf{y}\|_2 = \sqrt{\sum_{i=1}^{n} (x^{(i)} - y^{(i)})^2}.$$

Every prototype-based clustering algorithm is based on an objective function J that quantifies the goodness of the cluster model. J must be minimized to obtain optimal clusters. The algorithms determine the best decomposition by minimizing J.

The simplest algorithm is called hard c-means or *k-means* clustering. Here, each data point $\mathbf{x}_j$ in dataset $\mathcal{X} = \{\mathbf{x}_1,\dots,\mathbf{x}_m\}, \mathcal{X} \subseteq \mathrm{IR}^n$ is assigned to exactly one cluster $\Gamma_i \subset \mathcal{X}$. The set of clusters $\Gamma = \{\Gamma_1,\dots,\Gamma_c\}$ must be an exhaustive partition of $\mathcal{X}$ into c non-empty and pairwise disjoint subsets Γ_i, $1 < i < c$. The data partition is optimal when the sum of squared distances between cluster centers and data points assigned to them is minimal. The clusters should be as homogeneous as possible. The objective function of hard c-means is thus

$$J_h(X,U_h,C) = \sum_{i=1}^{c}\sum_{j=1}^{m} u_{ij} d_{ij}^2 \tag{16.12}$$

whereas d_{ij} is the distance between $\mathbf{c}_i$ and $\mathbf{x}_j$, $U = u_{ij} \in \{0, 1\}_{c\times m}$ is called *partition matrix* with

$$u_{ij} = \begin{cases} 1 & \text{if } \mathbf{x}_j \in \Gamma_i, \\ 0 & \text{otherwise.} \end{cases}$$

Equation 16.12 is minimized subject to the following two constraints: Each data point is assigned exactly to one cluster, i.e.,

$$\sum_{i=1}^{c} u_{ij} = 1, \quad \forall j \in \{1,\dots,m\}. \tag{16.13}$$

Every cluster must contain at least one data point, i.e.,

$$\sum_{j=1}^{m} u_{ij} > 0, \quad \forall i \in \{1,\dots,c\}. \tag{16.14}$$

J_h depends on both c and the assignment U of data points to the clusters. Finding the parameters that minimize J_h is NP-hard. Therefore J_h is minimized by *alternating optimization* (AO). The parameters to optimize are split into two groups. One group is optimized holding the other group fixed (and vice versa). An iterative update scheme is repeated until the algorithm converges. It cannot be guaranteed that a global optimum will be reached. Hence, the algorithm may get stuck in a local minimum. The AO scheme for hard c-means first choses c initial $\mathbf{c}_i$, e.g., by randomly picking c data points from $\mathcal{X}$. Then, C is fixed and U is determined that minimizes J_h. This is done by assigning each data point to its closest cluster center, i.e.,

$$u_{ij} = \begin{cases} 1 & \text{if } i = \arg\min_{k=1}^{c} d_{kj}, \\ 0 & \text{otherwise.} \end{cases}$$

After that U is fixed and $\mathbf{c}_i$ are updated as the mean of all $\mathbf{x}_j$ assigned to them. The mean minimizes the sum of square distances in J_h, i.e.,

$$\mathbf{c}_i = \frac{\sum_{j=1}^{m} u_{ij}\mathbf{x}_j}{\sum_{j=1}^{m} u_{ij}}.$$

Finally, both steps are repeated until no change in C or U can be observed.

The hard c-means algorithm tends to get stuck in local minimum. It is therefore necessary to conduct several runs with different initializations (Duda and Hart 1973). The best result of many clusterings can be chosen based on the value of J_h. The crisp memberships $u_{ij} \in \{0, 1\}$ prohibit ambiguous assignments. When clusters are badly delineated or overlapping, relaxing this requirement is needed. This can be achieved using fuzzy clustering.

Fuzzy clustering algorithms allow gradual memberships of data points to a cluster in $[0, 1]$. A data point can thus belong to more than one cluster. Consequently, the membership degrees offer finer degrees of detail and express how ambiguously $\mathbf{x}_j$ should belong to Γ_i. The clusters Γ_i have been classical subsets so far. Now, they are represented by fuzzy sets μ_{Γ_i} of $\mathcal{X}$. Instantly, the cluster assignment u_{ij} is the membership degree of $\mathbf{x}_j$ to Γ_i s.t. $u_{ij} = \mu_{\Gamma_i}(\mathbf{x}_j) \in [0, 1]$. Thence, a fuzzy label vector $\mathbf{u} = (u_{1j}, \ldots, u_{cj})^T$ is linked to each $\mathbf{x}_j$. The matrix $U = (u_{ij}) = (\mathbf{u}_1, \ldots, \mathbf{u}_m)$ is then called *fuzzy partition matrix*. Two types of fuzzy cluster partitions are known, i.e., *probabilistic* and possibilistic. The differ in the constraints they place on the membership degrees. For a *probabilistic cluster partition*, the constraints expressed by (16.13) and (16.14) must hold. So, no cluster can contain the full membership of all data points. Also, the membership degrees for a given datum resemble the probabilities of being member of a corresponding cluster. A *possibilistic cluster partition* only needs to fulfill the constraint (16.13). Here, we only focus on the former type of cluster partition. Algorithms based on the latter one can be found in Höppner et al. (1999).

In order to handle fuzzy membership assignments, we must minimize the objective function

$$J_f(X, U_h, C) = \sum_{i=1}^{c} \sum_{j=1}^{m} u_{ij}^{w} d_{ij}^{2}$$

subject to (16.13) and (16.14). The parameter $w \in \mathrm{IR}$ with $w > 1$ is called *fuzzifier*. The value of w determines the "fuzziness" of the grouping. For $w = 1$ (i.e., $J_h = J_f$), the assignments remain hard. Only fuzzifiers $w > 1$ lead to fuzzy memberships (Bezdek 1973). Thus the clusters become softer/harder with higher/lower w. Usually w is set to 2 in most applications. The function J_f is alternately optimized, i.e., first optimizing U for fixed cluster parameters $U_\tau = j_U(C_{\tau-1})$, then optimizing C for fixed membership degrees $C_\tau = j_C(U_\tau)$. The update formulas can be determined by setting the derivative of J_f w.r.t. U and C to zero. The resulting equations form the *fuzzy c-means* (FCM) algorithm. The membership degrees are chosen according to Bezdek (1981)

$$u_{ij} = \frac{1}{\sum_{k=1}^{c} \left(\frac{d_{ij}^2}{d_{kj}^2} \right)^{\frac{1}{w-1}}} = \frac{d_{ij}^{\frac{2}{1-w}}}{\sum_{k=1}^{c} d_{kj}^{\frac{2}{1-w}}}$$

which is independent of the chosen distance measure. For the basic FCM model With the second step of the AO scheme, the derivations of J_f w.r.t. the centers yield (Bezdek 1981)

$$\mathbf{c}_i = \frac{\sum_{j=1}^{m} u_{ij}^m \mathbf{x}_j}{\sum_{j=1}^{m} u_{ij}^m}.$$

Like hard c-means, FCM can be initialized with randomly placed cluster centers. Updating in the AO scheme can be stopped if the number of iterations τ exceeds some predefined $\tau_{\max}$ or if changes in the prototypes are smaller than some termination accuracy. FCM is stable and robust. Compared to hard c-means, it is quite insensitive to the initialization and not likely to get stuck in a local minimum. FCM converges in a saddle point or minimum (but not in a maximum) Bezdek (1981). Further fuzzy clustering algorithms, distance functions variants and applications can be found in Bezdek et al. (1999) and Höppner et al. (1999).

Fuzzy Rule Generation

The automatic generation of linguistic rules plays an important role in many applications, e.g., classification (Kuncheva 2000; Nauck and Kruse 1997), regression (Dickerson and Kosko 1996; Nauck and Kruse 1999; Wang and Mendel 1992), control engineering (Klawonn et al. 1995; Klawonn and Kruse 1993, 1995, 1997), image processing (Bezdek et al. 1999; Höppner et al. 1999). In fuzzy data analysis, we are interested in learning fuzzy rules from observations using fuzzy methods, e.g., FCM.

Before we talk about the generation of linguistic rules from fuzzy clustering, let us briefly mention the some other methods based on fuzzy logic. Grid-based approaches define fixed fuzzy partitions for every variable. Every cell in that multidimensional grid may correspond to one rule (Wang and Mendel 1992). Most well-known are hybrid methods to induce fuzzy rules. Therefore a fuzzy system is combined with computational intelligence techniques. For instance, *evolutionary algorithms* are used for guided searching the space of possible rule bases (Cordón et al. 2004). *Neuro-fuzzy systems* use learning methods of artificial neural network (e.g., backpropagation) to tune parameters of a network that can be directly understood as a fuzzy system (Nauck et al. 1997). Standard rule generation methods have been fuzzified as well (e.g., separate-and-conquer rule learning (Hühn and Hüllermeier 2009), decision trees (Olaru and Wehenkel 2003), support vector machines (Moewes and Kruse 2008).

Here, we will restrict ourselves to FCM for fuzzy rule generation. Consider again the input space $X \subset \mathrm{IR}^n$ and the output space $Y \subset \mathrm{IR}$. We observe m patterns $(\mathbf{x}_j, y_j) \in S \subseteq X \times Y$ where $j = 1, \ldots, m$. Running FCM on that dataset S leads to c cluster prototypes $\mathbf{c}_i = (c_i^{(1)}, \ldots, c_i^{(n)}, c_i^{(y)})$ with $i = 1, \ldots, c$ that can be seen as concatenation of both the input values $c_i^{(j)}, j = 1, \ldots, n$ and the output value $c_i^{(y)}$. Thus every prototype represents one linguistic rule

$$R_i : \text{ if } x \text{ is close to } \left(c_i^{(1)}, \ldots, c_i^{(n)}\right) \text{ then } y \text{ is close to } c_i^{(y)}.$$

Using the membership degrees U, we can rewrite these rules as

$$R_i : \text{ if } \mathbf{u}_i^{\mathbf{x}}(\mathbf{x}) \text{ then } u_i^y(y). \tag{16.15}$$

The only problem is that FCM returns the membership degrees $\mathbf{u}_i(\mathbf{x}, y)$ of the product space $X \times Y$. To obtain rules like (16.15), we must *project* $\mathbf{u}_i$ onto $\mathbf{u}_i^{\mathbf{x}}$ and u_i^y. If $\mathbf{x}$ and y are restricted to $[\mathbf{x}_{\min}, \mathbf{x}_{\max}]$ and $[y_{\min}, y_{\max}]$, respectively, the projections are given by

$$\mathbf{u}_i^{\mathbf{x}}(\mathbf{x}) = \sup_{y \in [y_{\min}, y_{\max}]} \mathbf{u}_i(\mathbf{x}, y),$$
$$u_i^y(y) = \sup_{x \in [\mathbf{x}_{\min}, \mathbf{y}_{\max}]} \mathbf{u}_i(\mathbf{x}, y).$$

We can also project $\mathbf{u}_i$ onto each single input variable $X_1, \ldots, X_n$ by

$$u_{ik}(x^{(k)}) = \sup_{\mathbf{x}^{(\neg k)} \in [\mathbf{x}_{\min}^{(\neg k)}, \mathbf{x}_{\max}^{(\neg k)}]} \mathbf{u}_i^{\mathbf{x}}(\mathbf{x})$$

for $k = 1, \ldots, n$ where as $\mathbf{x}^{(\neg k)} \stackrel{\text{def}}{=} (x^{(1)}, \ldots, x^{(k-1)}, x^{(k+1)}, \ldots, x^{(n)})$. We may thus write (16.15) in form of a *Mamdani-Assilian rule* (16.6) as

$$R_i : \text{ if } \bigwedge_{k=1}^{n} u_{ik}(x^{(k)}) \text{ then } u_i^y(y). \tag{16.16}$$

For one rule, the output value of an unseen input $\mathbf{x} \in \mathrm{IR}^n$ will be equivalent to (16.7) if the minimum t-norm is used as conjunction $\wedge$. The overall output of the complete rule base is given by a disjunction $\vee$ of all rule outputs (cf. (16.9) if $\vee$ is the t-conorm maximum).

A crisp output can then again be computed by defuzzification, e.g., using the COG method (16.10). Since this computation is rather costly, the output membership functions u_i^y are commonly be replaced by singletons, i.e.,

$$u_i^y(y) = \begin{cases} 1 & \text{if } y = c_i^{(y)}, \\ 0 & \text{otherwise.} \end{cases}$$

Since each rule consequent comprise the component $c_i^{(y)}$ of the cluster prototype, we can rewrite (16.16) as *Sugeno-Yasukawa rule* (Sugeno and Yasukawa 1993)

$$R_i: \text{ if } \bigwedge_{k=1}^{n} u_{ik}(x^{(k)}) \text{ then } y = c_i^{(y)}.$$

These rules strongly resemble the neurons of an RBF network. This will become clear if every membership function is Gaussian, i.e.,

$$\mathbf{u}_i^{\mathbf{x}}(\mathbf{x}) = \exp\left(\frac{\mathbf{x} - \mu_i}{\sigma_i}\right)^2,$$

and if there are normalized, i.e., $\sum_{i=1}^{c} \mathbf{u}_i^{\mathbf{x}}(\mathbf{x}) = 1$ for all $\mathbf{x} \in \mathrm{IR}^n$. This link is used in neuro-fuzzy systems for both training fuzzy rules with backpropagation and initializing RBF networks with fuzzy rules (Nauck and Kruse 1997).

Transfer Passenger Analysis Based on FCM

The German Aerospace Center (DLR) developed a macroscopic passenger flow model for simulating passenger movements on airport's land side. For the passenger movements in terminal areas, probabilistic distribution functions are used today. In (Keller and Kruse 2002), the goal was to build a fuzzy rule base describing the transfer passenger amount between aircrafts. These rules could be used to improve the macroscopic simulation. The key idea was to find the rules based on FCM. The following attributes of passengers were used to for analysis:

- The maximal amount of passengers in a certain aircraft (depending on the type of the aircraft)
- The distance between the airport of departure and the airport of destination (in three categories: short-, medium-, and long-haul)
- The time of departure
- The percentage of transfer passengers in the aircraft

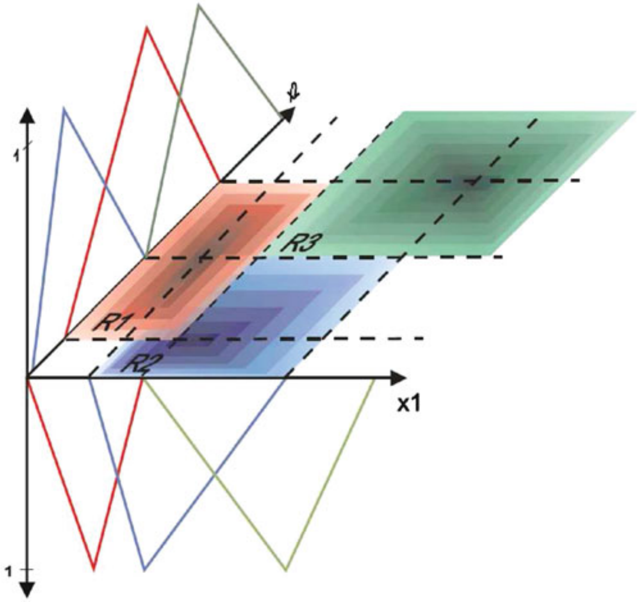


Fig. 16.10 Fuzzy rules and induced vague areas

The number of clusters were determined by validity measures (Höppner et al. 1999; Kruse et al. 2007) evaluating the whole partition of all data. The clustering was run for a varying number of clusters. The validity of the resulting partitions was compared based on the used measures.

An example of resulting fuzzy clusters are shown in Fig. 16.10. Every fuzzy cluster corresponds to one fuzzy rule. The color intensity indicates the firing strength of a specific rule. The vague areas are the fuzzy clusters whereas the color intensity indicates the membership degree. The tips of the fuzzy partitions are obtained in every domain by projections of the multidimensional cluster centers (as explained before in section "Fuzzy Rule Generation").

The fuzzy rules obtained by FCM were simplified through several steps. First, similar fuzzy sets were combined to one fuzzy set. Fuzzy sets similar to the universal fuzzy set were removed. Fuzzy rules with the same input clauses were either combined if they also shared the same output clauses or else they were removed from the rule base. Finally, around five rules could be obtained from FCM. Among them were the two following rules: If an aircraft with a relatively small amount of maximal passengers (80–200) has a short- or medium-haul destination departing late at night, then usually this flight has a high amount of transfer passengers (80–90%). If a flight with a medium-haul destination and a small aircraft (about 150 passengers) starts about noon, then it carries a relatively high amount of transfer passengers (ca. 70%). We refer to Keller and Kruse (2002) for more details about this real-world application.

Acknowledgment R. Belohlavek was supported by the ESF project No. CZ.1.07/2.3.00/20.0059 (co-financed by the European Social Fund and the state budget of the Czech Republic).

References

Bezdek, J.C.: Fuzzy mathematics in pattern classification. PhD thesis, Cornell University, Itheca, NY, USA (1973)

Bezdek, J.C.: Pattern Recognition with Fuzzy Objective Function Algorithms. Kluwer Academic Publishers, Norwell, MA, USA (1981)

Bezdek, J.C., Keller, J., Krisnapuram, R., Pal, N.R.: Fuzzy Models and Algorithms for Pattern Recognition and Image Processing, *The Handbooks of Fuzzy Sets*, vol. 4. Kluwer Academic Publishers (1999)

Boixader, D., Jacas, J.: Extensionality based approximate reasoning. International Journal of Approximate Reasoning **19**(3–4), 221–230 (1998). DOI 10.1016/S0888-613X(98)00018-8

Belohlavek, R.: Fuzzy Relational Systems: Foundations and Principles, *IFSR International Series on Systems Science and Engineering*, vol. 20. Kluwer Academic Publishers (2002)

Belohlavek, R., Klir, G.J.: On Elkan's theorems: Clarifying their meaning via simple proofs: Research articles. International Journal of Intelligent Systems **22**, 203–207 (2007). DOI 10.1002/int.v22:2. ACM ID: 1190438

Belohlavek, R., Klir, G.J. (eds.): Concepts and Fuzzy Logic. MIT Press, Cambridge, MA, USA (2011)

Belohlavek, R., Vychodil, V.: Fuzzy Equational Logic, *Studies in Fuzziness and Soft Computing*, vol. 186. Springer (2005)

Belohlavek, R., Vychodil, V.: Attribute implications in a fuzzy setting. In: Formal Concept Analysis, *Lecture Notes in Computer Science*, vol. 3874, pp. 45–60. Springer-Verlag (2006)

Cheeseman, P.: An inquiry into computer understanding. Computational Intelligence **4**, 58–66 (1988)

Cheeseman, P.: Probabilistic versus fuzzy reasoning. In: L.N.K. Kanal, Lemmer (eds.) UAI '85: Proceedings of the First Annual Conference on Uncertainty in Artificial Intelligence. Elsevier, New York, NY, USA (1988)

Cignoli, R., Esteva, F., Godo, L., Torrens, A.: Basic fuzzy logic is the logic of continuous t-norms and their residua. Soft Computing **4**(2), 106–112 (2000). DOI 10.1007/s005000000044

Cordón, O., del Jesus, M.J., Herrera, F.: A proposal on reasoning methods in fuzzy rule-based classification systems. International Journal of Approximate Reasoning **20**(1), 21–45 (1999). DOI 10.1016/ S0888-613X(00)88942-2

Cordón, O., Gomide, F., Herrera, F., Hoffmann, F., Magdalena, L.: Ten years of genetic fuzzy systems: current framework and new trends. Fuzzy Sets and Systems **141**(1), 5–31 (2004). DOI 10.1016/S0165-0114(03) 00111-8

Dickerson, J.A., Kosko, B.: Fuzzy function approximation with ellipsoidal rules. IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics **26**(4), 542–560 (1996). DOI 10.1109/3477.517030

Dubois, D., Prade, H.: The generalized modus ponens under sup-min composition – a theoretical study. In: M.M. Gupta, A. Kandel, W. Bandler, J.B. Kiszka (eds.) Approximate Reasoning in Expert Systems, pp. 217–232. Elsevier Science Publisher B.V. (North-Holland), Amsterdam, Netherlands (1985)

Dubois, D., Prade, H.: Possibility Theory: An Approach to Computerized Processing of Uncertainty. Plenum Press, New York, NY, USA (1988)

Dubois, D., Prade, H.: Possibility theory as a basis for preference propagation in automated reasoning. In: 1992 IEEE International Conference on Fuzzy Systems, pp. 821–832. IEEE Press, New York, NY, USA (1992). DOI 10. 1109/FUZZY.1992.258765

Duda, R.O., Hart, P.E.: Pattern Classification and Scene Analysis. John Wiley & Sons, Ltd., New York, NY, USA (1973)

Elkan, C.: The paradoxical success of fuzzy logic. In: Proceedings of the 11th National Conference on Artificial Intelligence, pp. 698–703. AAAIPress/MIT Press, Cambridge, MA, USA (1993)

Elkan, C.: The paradoxical success of fuzzy logic. IEEE Expert: Intelligent Systems and Their Applications **9**, 3–8 (1994). DOI 10.1109/64.336150. ACM ID: 630036

Gerla, G.: Fuzzy Logic: Mathematical Tools for Approximate Reasoning. Kluwer Academic Publishers, Dordrecht, Netherlands (2001)

Goguen, J.A.: The logic of inexact concepts. Synthese **19**(3–4), 325–373 (1968). DOI 10.1007/BF00485654

Gottwald, S.: A Treatise on Many-Valued Logics, *Studies in Logic and Computation*, vol. 9. Research Studies Press, Baldock, Hertfordshire, England (2001)

Hájek, P.: Metamathematics of Fuzzy Logic, *Trends in Logic*, vol. 4. Kluwer Academic Publishers, Boston, MA, USA (1998)

Hájek, P.: What is mathematical fuzzy logic. Fuzzy Sets and Systems **157**(5), 597–603 (2006). DOI 10.1016/j.fss.2005.10.004

Höppner, F., Klawonn, F., Kruse, R., Runkler, T.: Fuzzy Cluster Analysis: Methods for Classification, Data Analysis and Image Recognition. John Wiley & Sons, Ltd., New York, NY, USA (1999)

Hühn, J.C., Hüllermeier, E.: FR3: a fuzzy rule learner for inducing reliable classifiers. IEEE Transactions on Fuzzy Systems **17**(1), 138–149 (2009). DOI 10.1109/TFUZZ.2008.2005490

Keller, A., Kruse, R.: Fuzzy rule generation for transfer passenger analysis. In: L. Wang, S.K. Halgamuge, X. Yao (eds.) Proceedings of the 1st International Conference on Fuzzy Systems and Knowledge Discovery (FSDK'02), pp. 667–671. Orchid Country Club, Singapore (2002)

Klawonn, F., Castro, J.L.: Similarity in fuzzy reasoning. Mathware & Soft Computing **2**(3), 197–228 (1995)

Klawonn, F., Gebhardt, J., Kruse, R.: Fuzzy control on the basis of equality relations with an example from idle speed control. IEEE Transactions on Fuzzy Systems **3**(3), 336–350 (1995). DOI 10.1109/91.413237

Klawonn, F., Kruse, R.: Equality relations as a basis for fuzzy control. Fuzzy Sets and Systems **54**(2), 147–156 (1993). DOI 10.1016/0165-0114(93) 90272-J

Klawonn, F., Kruse, R.: Automatic generation of fuzzy controllers by fuzzy clustering. In: 1995 IEEE International Conference on Systems, Man, and Cybernetics: Intelligent Systems for the 21st Century, vol. 3, pp. 2040–2045. IEEE Press, Vancouver, BC, Canada (1995). DOI 10.1109/ ICSMC.1995.538079

Klawonn, F., Kruse, R.: Constructing a fuzzy controller from data. Fuzzy Sets and Systems **85**(2), 177–193 (1997). DOI 10.1016/0165-0114(95)00350-9

Klawonn, F., Novák, V.: The relation between inference and interpolation in the framework of fuzzy systems. Fuzzy Sets and Systems **81**(3), 331–354 (1996). DOI 10.1016/0165-0114(96) 83710-9

Klement, E.P., Mesiar, R., Pap, E.: Triangular Norms, *Trends in Logic*, vol. 8. Kluwer Academic Publishers, Dordrecht, Netherlands (2000)

Klir, G.J.: Is there more to uncertainty than some probability theorists might have us belive? International Journal of General Systems **15**(4), 347–378 (1989). DOI 10.1080/03081078908935057

Klir, G.J., Yuan, B.: Fuzzy Sets and Fuzzy Logic: Theory and Applications. Prentice Hall, Upper Saddle River, NJ, USA (1995)

Kosko, B.: Fuzziness vs. probability. International Journal of General Systems **17**(2), 211–240 (1990). DOI 10.1080/03081079008935108

Kruse, R., Döring, C., Lesot, M.: Fundamentals of fuzzy clustering. In: J.V. de Oliveira, W. Pedrycz (eds.) Advances in Fuzzy Clustering and its Applications, pp. 3–30. John Wiley & Sons, Ltd., Chichester, UK (2007)

Kruse, R., Gebhardt, J., Klawonn, F.: Foundations of Fuzzy Systems. John Wiley & Sons Ltd, Chichester, UK (1994)

Kuncheva, L.I.: Fuzzy Classifier Design, *Studies in Fuzziness and Soft Computing*, vol. 49. Physica-Verlag, Heidelberg, New York (2000)

Lindley, D.V.: The probability approach to the treatment of uncertainty in artificial intelligence and expert systems. Statistical Science **2**(1), 17–24 (1987)

Loginov, V.I.: Probability treatment of zadeh membership functions and their use in pattern recognition. Engineering Cybernetics **4**(2), 68–69 (1966)

Mamdani, E.H., Assilian, S.: An experiment in linguistic synthesis with a fuzzy logic controller. International Journal of Man-Machine Studies **7**(1), 1–13 (1975). DOI 10.1016/S0020-7373 (75)80002-2

Moewes, C., Kruse, R.: Unification of fuzzy SVMs and rule extraction methods through imprecise domain knowledge. In: J.L. Verdegay, L. Magdalena, M. Ojeda-Aciego (eds.) Proceedings of the International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU-08), pp. 1527–1534. Torremolinos (Málaga) (2008)

Nauck, D., Klawonn, F., Kruse, R.: Foundations of Neuro-Fuzzy Systems. John Wiley & Sons, Inc. (1997)

Nauck, D., Kruse, R.: A neuro-fuzzy method to learn fuzzy classification rules from data. Fuzzy Sets and Systems **89**(3), 277–288 (1997). DOI 10.1016/ S0165-0114(97)00009-2

Nauck, D., Kruse, R.: Neuro-fuzzy systems for function approximation. Fuzzy Sets and Systems **101**(2), 261–271 (1999). DOI 10.1016/S0165-0114(98) 00169-9

Novák, V., Perfilieva, I., Močkoř, J.: Mathematical Principles of Fuzzy Logic. Kluwer Academic Publishers, Dordrecht, Netherlands (1999)

Olaru, C., Wehenkel, L.: A complete fuzzy decision tree technique. Fuzzy Sets and Systems **138** (2), 221–254 (2003). DOI 10.1016/S0165-0114(03) 00089-7

Pavelka, J.: On fuzzy logic i, II, III. Zeitschrift für Mathematische Logik und Grundlagen der Mathematik **25**, 45–52, 119–134, 447–464 (1979)

Schröder, M., Petersen, R., Klawonn, F., Kruse, R.: Two paradigms of automotive fuzzy logic applications. In: M. Jamshidi, A. Titli, L. Zadeh, S. Boverie (eds.) Applications of Fuzzy Logic: Towards High Machine Intelligence Quotient Systems, *Environmental and Intelligent Manufacturing Systems Series*, vol. 9, pp. 153–174. Prentice-Hall, Inc., Upper Saddle River, NJ, USA (1997)

Smith, N.J.J.: Vagueness and Degrees of Truth. Oxford University Press, New York, NY, USA (2009)

Stevens, S.S.: On the theory of scales of measurement. Science **103**(2684), 677–680 (1946). DOI 10.1126/science.103.2684.677

Sudkamp, T.: Similarity, interpolation, and fuzzy rule construction. Fuzzy Sets and Systems **58**(1), 73–86 (1993). DOI 10.1016/0165-0114(93) 90323-A

Sugeno, M., Yasukawa, T.: A fuzzy-logic-based approach to qualitative modeling. IEEE Transactions on Fuzzy Systems **1**(1), 7–31 (1993). DOI 10. 1109/TFUZZ.1993.390281

Takagi, T., Sugeno, M.: Fuzzy identification of systems and its applications to modeling and control. IEEE Transactions on Systems, Man, and Cybernetics **15**(1), 116–132 (1985)

van Deemter, K.: Not Exactly: In Praise of Vagueness. Oxford University Press, New York, NY, USA (2010)

Wang, L., Mendel, J.M.: Generating fuzzy rules by learning from examples. IEEE Transactions on Systems, Man, and Cybernetics **22**(6), 1414–1427 (1992). DOI 10.1109/21.199466

Zadeh, L.A.: Fuzzy sets. Information and Control **8**(3), 338–353 (1965). DOI 10.1016/S0019-9958 (65)90241-X

Zadeh, L.A.: Probability measures of fuzzy events. Journal of Mathematical Analysis and Applications **23**(2), 421–427 (1968). DOI 10.1016/ 0022-247X(68)90078-4

Zadeh, L.A.: Outline of a new approach to the analysis of complex systems and decision processes. IEEE Transactions on Systems, Man, and Cybernetics **3**(1), 28–44 (1973)

Zadeh, L.A.: Fuzzy logic and approximate reasoning. Synthese **30**(3-4), 407–428 (1975). DOI 10.1007/BF00485052

Zadeh, L.A.: Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets and Systems **1**(1), 3–28 (1978). DOI 10.1016/0165-0114(78)90029-5. (Reprinted in Fuzzy Sets and Systems 100 (Supplement 1), 9-34, (1999))

Zadeh, L.A.: A theory of approximate reasoning. In: J.E. Hayes, D. Michie, L.I. Mikulich (eds.) Proceedings of the Ninth Machine Intelligence Workshop, no. 9 in Machine Intelligence, pp. 149–194. John Wiley & Sons, Ltd., New York, NY, USA (1979)

Zadeh, L.A.: The role of fuzzy logic in the management of uncertainty in expert systems. Fuzzy Sets and Systems **11**(1–3), 197–198 (1983). DOI 10.1016/S0165-0114(83)80081-5

Zadeh, L.A.: Toward a perception-based theory of probabilistic reasoning with imprecise probabilities. Journal of Statistical Planning and Inference **105**(1), 233–264 (2002). DOI 10.1016/S0378-3758(01)00212-9

Zadeh, L.A.: Generalized theory of uncertainty (GTU)-principal concepts and ideas. Computational Statistics & Data Analysis **51**(1), 15–46 (2006). DOI 10.1016/j.csda.2006.04.029

Zadeh, L.: Toward human level machine intelligence - is it achievable? the need for a paradigm shift. IEEE Computational Intelligence Magazine **3**(3), 11–22 (2008). DOI 10.1109/MCI.2008.926583