Fuzzy Mathematics

- Fuzzy concepts are 'natural' generalizations of conventional mathematical concepts.
- Probability and possibility are complementary.
- Possibility theory is the attempt to be precise about uncertainty, to related statistical objects with rule-based and fuzzy concepts.

Everything is a matter of degree.

—The Fuzzy Principle

Since almost all objects of mathematics can be described by sets (e.g., a function as a set of ordered pairs), one can establish fuzzy generalizations for nearly all of those. In the previous sections, we have seen that fuzzy mathematical concepts occur 'naturally', that is, considering the uncertainty involved in the application, the need to generalize creates fuzzy mathematical objects. We considered

- fuzzy sets
- fuzzy mappings
- fuzzy relations and their composition
- fuzzy graphs
- fuzzy clustering
- possibility measures and distributions

To apply formal mathematical objects to an engineering problem, mathematical statements are formulated using propositional calculus for which the standard logical connectives and their set-theoretic equivalents for propositions $p$ and $q$ are:
Conjunction $p \land q$ “$p$ and $q$” $A_p \cap B_q$ Intersection
Disjunction $p \lor q$ “$p$ or $q$” $A_p \cup B_q$ Union
Negation $\neg p$ “not $p$” $A_p^c$ Complement
Implication $p \implies q$ “$p$ implies $q$” $A_p^c \cup B_q$ Entailment

Classically, the subsets $A$ of a set $X$ form a Boolean algebra under the operations union, intersection, and complement. In particular, the double complement of a subset $A$ is $A$ itself, $(A^c)^c = A$, in parallel to the tautology $\neg \neg p = p$ of the propositional calculus. Fuzzy versions of these operations above will abandon the law of the excluded middle, $A \cap A^c = \emptyset$, and consequently allow us to model vague, fuzzy, and ambiguous concepts, avoiding paradoxes. From these extensions, we find that there are at least five ways to develop rule-based systems:

1. Composite fuzzy relations (approximate reasoning),
2. Functional approximation (e.g., Takagi-Sugeno models),
3. Similarity-based reasoning,
4. Multi-valued logic,
5. Possibilistic reasoning,

of which the first three were addressed in this book. An overview of different fuzzy mathematical extensions to classical concepts is given in Figure 11.1. In this book, fuzzy sets have been described in different ways. Most commonly a fuzzy set is considered to be a family of pairs, $R = \{(x, \mu_R(x))\}$ and membership function $\mu(\cdot)$. $\mu_R(x)$ describes the degree of membership of $x$ in fuzzy set $R$. Viewing degrees of membership as some kind of weighting on elements of the underlying reference space $X$, a fuzzy restriction is the mapping $\mu$

$$
\begin{align*}
\mu : & \quad X \rightarrow [0, 1] \\
& \quad x \mapsto \mu(x). 
\end{align*}
$$

Especially in engineering, triangular and trapezoidal fuzzy set-membership functions are sufficiently accurate and have the advantage of a simple implementation. The following equations are useful in this context. Let $a \leq b \leq c \leq d$ denote characteristic points:

“Left-open” set: $\mu(x; a, b) = \max \left( \min \left( \frac{b-x}{b-a}, 1 \right), 0 \right)$

“Right-open”: $\mu(x; a, b) = \max \left( \min \left( \frac{x-a}{b-a}, 1 \right), 0 \right)$

“Triangular”: $\mu(x; a, b, c) = \max \left( \min \left( \frac{x-a}{b-a}, \frac{c-x}{c-b}, 0 \right) \right)$

“Trapezoidal”: $\mu(x; a, b, c, d) = \max \left( \min \left( \frac{x-a}{b-a}, 1, \frac{d-x}{d-c}, 0 \right) \right)$
Instead of taking elements of the universe of discourse as arguments, we may consider the co-domain of \( \mu \) to describe subsets of \( X \) in terms of \( \alpha \)-cuts, \( R^\alpha \) where \( R^\alpha = \{ x : \mu_R(x) \geq \alpha \} \) is also called level-set. The fact that a family of level-sets can describe a fuzzy set is manifested in the decomposition or representation theorem. Instead of taking a set-membership perspective we may view \( \mu(\cdot) \) as a mapping, or fuzzy restriction \( R \):

\[
\mu_R : X \rightarrow L \\
x \mapsto \alpha,
\]

where here we assume \( L = [0, 1] \). While in the set-membership setting we first identify a value \( x \) and then determine its degree of membership, we may also start with a level \( \alpha \in L \) to find out which elements in \( X \) satisfy this condition. This leads to the definition of a level-set or \( \alpha \)-cut \( R_\alpha \) (see Figure 11.2):

\[
R_\alpha = \{ x \in X : \mu(x) \geq \alpha \} . \quad (11.2)
\]

The representation theorem [NW97] shows that a family of sets \( \{ R^\alpha \} \) with the assertion “\( x \) is in \( R^\alpha \)” has the “degree of truth” \( \alpha \), and composes a fuzzy set, or equivalently, a fuzzy restriction

\[
R(x) = \sup_{\alpha \in (0,1]} \min(\alpha, \zeta_{R^\alpha}(x)) , \quad (11.3)
\]

where

\[
\alpha = \min\{\alpha, \zeta_{R^\alpha}(x)\}
\]

and hence

\[
R = \{ (x, \alpha) : x \in X, \mu_R(x) = R(x) = \alpha \} . \quad (11.4)
\]
We can summarize the level-set representation of \( R \) as the mapping

\[
R : \left\{ (R_\alpha) \right\} \rightarrow \mathcal{F}(X)
\]

\( (R_\alpha) \mapsto \mu_R \),

where

\[
\mu_R(x) = \sup_\alpha \{ \alpha \in [0, 1], \ x \in R_\alpha \} .
\]

**Fig. 11.2** Level-set representation of fuzzy restriction (set) \( R \).

Figure 11.2 suggests yet another perspective on fuzzy restrictions as multi-valued maps. Let \( \Gamma \) be a map from \( L \) to \( X \) such that for any \( \alpha \in L \), the image is made up of those \( x \in X \) which are compatible with \( \alpha \):

\[
\Gamma : \ L \rightarrow \mathcal{P}(X) \quad (11.5)
\]

\( \alpha \mapsto \Gamma(\alpha) \).

Expressed in terms of ordered pairs \( (\alpha, x) \in L \times X \),

\[
R = \{(\alpha, x) : x \in \Gamma(\alpha)\} , \quad (11.6)
\]

which is called *compatibility relation* and is related to \( \Gamma(\alpha) \) by

\[
\Gamma(\alpha) = \{x : (\alpha, x) \in R\} . \quad (11.7)
\]

We also considered generalizations of ordinary relations to fuzzy equivalents. For example, considering the binary relation

\[
R : X \times Y \rightarrow \{0, 1\}
\]

\( (x, y) \mapsto \zeta_R(x, y) \),

a fuzzy relation \( R \) is a fuzzy set defined in \( X \times Y \) as

\[
R : X \times Y \rightarrow [0, 1]
\]

\( (x, y) \mapsto \mu_R(x, y) \).
11.1 THE ALGEBRA OF FUZZY SETS

In this section, we first summarize the usual use and notation of (fuzzy) sets and then look at more formal aspects of how to describe sets.

For 'crisp' sets, if $X$ is any set and $x \in X$, the algebra of the power set $\mathcal{P}(X)$ of $X$, that is, of the set of (crisp) subsets of $X$, is usually formulated in terms of $A \in \mathcal{P}(X)$ and $B \in \mathcal{P}(X)$ as follows:

- Containment: $A \subseteq B \iff x \in A \Rightarrow x \in B$
- Equality: $A = B \iff A \subseteq B$ and $B \subseteq A$
- Complement: $A^c = \{x \in X : x \notin A\}$
- Intersection: $A \cap B = \{x \in X : x \in A$ and $x \in B\}$
- Union: $A \cup B = \{x \in X : x \in A$ or $x \in B$ or both\}$.

For fuzzy sets realized via functions, the set-theoretic operations and relations above have their equivalents in $\mathcal{F}(X)$, namely, the set of all fuzzy sets (the set of all membership functions). Let $\mu_A$ and $\mu_B$ be in $\mathcal{F}(X)$:

- Containment: $A \subseteq B \iff \mu_A(x) \leq \mu_B(x)$
- Equality: $A = B \iff \mu_A(x) = \mu_B(x)$
- Complement: $\mu_A^c(x) = 1 - \mu_A(x)$
- Intersection: $\mu_{A \cap B}(x) = T(\mu_A(x), \mu_B(x))$
- Union: $\mu_{A \cup B}(x) = S(\mu_A(x), \mu_B(x))$.

These are by no means the only definitions but those most commonly used. Similarly, the intersection and union operators are most commonly defined by

\[ \mu_{A \cap B}(x) = \min(\mu_A(x), \mu_B(x)) \quad \text{and} \quad \mu_{A \cap B}(x) = \max(\mu_A(x), \mu_B(x)). \]

With this definition, the fuzzy union of $A$ and $B$ is the smallest fuzzy set containing both $A$ and $B$, and the intersection is the largest fuzzy set contained by both $A$ and $B$. This pair of operators is the only pair which preserves the equalities $\mu \cap \mu = \mu$ and $\mu \cup \mu = \mu$. In other words, it is the only pair of distributive and thus absorbing and idempotent pair of $t$-norm and $t$-conorm [FR94], that is, $(\mathcal{F}(X), \cap, \cup)$ is a distributive lattice. However, $(\mathcal{F}(X), \cap, \cup)$ is not a Boolean algebra, since not all elements of the lattice have complements in it. Considering possibility distributions on product spaces, using max, min, joint distributions are separable, that is, it is possible to project marginal distributions onto subspaces. Subsequently the concept of non-interactivity can be defined in analogy to independence in probability theory.

We also made frequent use of the fact that there are two distinct ways of formalizing the notion of a set. A subset $A$ of $X$ may be formulated as a list
describing its members

\[ A = \{ x \in X : x \text{ satisfies some condition} \} . \]

On the other hand, the subset \( A \) of \( X \) can be considered as an inclusion \( A \subset X \) and hence as a mapping. Equivalent subobjects of \( X \) are represented by the same “arrow” from \( X \), by taking the arrow to be the characteristic function

\[ \zeta_A : X \to \{0, 1\} \]

\[ x \mapsto \zeta_A(x) \quad \text{where} \quad \zeta_A(x) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \notin A. \end{cases} \]

Moreover, from the characteristic function \( \zeta \), one can reconstruct the subset \( A \): it consists of exactly those elements \( x \in X \) which land on 1 under \( \zeta \). In other words, \( A \) is the pullback\(^1\) of the inclusion \( \{1\} \subset \{0, 1\} \),

![Commutative diagram](image_url)

This can be generalized by replacing the set \( \{0, 1\} \) of ‘truth values’ by a non-negative interval \( L \) leading to ‘categorical’ formulations of fuzzy sets [Höh88]. This includes the case of \( L = [0, 1] \) being the unit-interval or some lattice defined on it. Then, the three operations \( \wedge, \vee, \text{ and } \neg \), defined by truth tables, are functions as \( L \times L \to L \).

### 11.2 THE EXTENSION PRINCIPLE

In Section 7.2, instead of viewing the fuzzy systems as an algorithm based on formal multi-valued logic, the rule-base and inference mechanism was described as a mapping from a fuzzy set \( A' \) in \( X \) to a fuzzy set \( B' \) in \( Y \). The compositional rule of inference generalized the ‘crisp’ rule

\[ \text{IF } x = a \text{ AND } y = f(x), \text{ THEN } y = f(a) \]

\(^1\)In the commutative diagram, \( A \) completes the solid line ‘corner’ by means of the dashed lines and is thus called pullback. The terminology originates from Category Theory [LS97] in which algebra is organized to look not just at the objects but also at the mappings (arrows) between them. Examples are functions between sets or continuous maps between spaces. A category then consists of objects and arrows between them. Category Theory subsequently starts with functions between sets rather than with sets and their elements.
The extension principle to be valid for fuzzy sets
\[ \mu_{B'}(y) = \sup_{x \in X} T(\mu_{A'}(x), \mu_R(x, y)) . \]  

(7.15)

The fuzzy system, defined by the compositional rule of inference, maps fuzzy sets in \( X \) to fuzzy sets in \( Y \). In other words, the fuzzy model describes a fuzzy mapping
\[ \tilde{f} : \mathcal{F}(X) \rightarrow \mathcal{F}(Y) \]
\[ \mu_A(x) \mapsto \tilde{f}(A) , \]

where we obtain \( \mu_{\tilde{f}(A)}(y) \) as a special case of the composition of two fuzzy relations
\[ \mu_{\tilde{f}(A)}(y) = \sup_{x \in X} T(\mu_{A_{\text{ext}}}(x, y), \mu_R(x, y)) \]  

(11.9)

with extension \( \mu_{A_{\text{ext}}}(x, y) = \mu_A(x) \), equivalent to (7.15) or the individual-rule-based inference (8.6).

We can take the extension of the mapping to the fuzzy mapping as a blueprint for a general extension principle. Let \( f \) be a mapping from \( X \) to \( Y \), \( y = f(x) \). Consider the situation where we are given a fuzzy number \( A \) ("approximately \( x_0 \)) instead of a real number. We wish to find the fuzzy image \( B \) by a generalization of \( f \); how do we construct \( B = f(A) \)? We would require that the membership values of \( B \) should be determined by the membership values of \( A \). Also \( \sup B \) should be the image of \( \sup A \) as defined by \( f \). If the function \( f \) is surjective (onto), that is not injective (not a one-to-one mapping), we need to choose which of the values \( \mu_A(x) \) to take for \( \mu_B(y) \). Zadeh proposed the sup-union of all values \( x \) with \( y = f(x) \) that have the membership degree \( \mu_A(x) \). In other words,
\[ \mu_B(y) = \sup_{x : y = f(x)} \mu_A(x) . \]  

(11.10)

In general, we have the mapping
\[ f : X_1 \times \cdots \times X_r \rightarrow Y \]
\[ (x_1, \ldots, x_r) \mapsto y = f(x_1, \ldots, x_r) , \]

which we aim to generalize to a function \( \tilde{f}(\cdot) \) of fuzzy sets. The extension principle is defined as
\[ \tilde{f} : \mathcal{F}(X) \rightarrow \mathcal{F}(Y) \]
\[ (\mu_{A_1}(x_1), \ldots, \mu_{A_r}(x_r)) \mapsto \tilde{f}(\mu_{A_1}(x_1), \ldots, \mu_{A_r}(x_r)) , \]

where
\[ \mu_{\tilde{f}(A_1, \ldots, A_r)}(y) = \sup_{(x_1, \ldots, x_r) \in f^{-1}(y)} \{ \mu_{A_1}(x_1) \wedge \cdots \wedge \mu_{A_r}(x_r) \} . \]  

(11.11)
In Section 1, we introduced the functional representation (model \( \mathcal{M} \)) of a system \( \mathcal{G} \), \( y = f(x) \) and its associated graph, \( F = \{(x, f(x))\} \). We then introduced various (equivalent) generalizations into fuzzy models described by the fuzzy graph \( \tilde{F} \): In Section 4.6, equation (4.33), in Section 5, equation (5.5), in Section 6, by (6.19), and in Section 8 by equation (8.5).

11.3 FUZZY RULES AND FUZZY GRAPHS

There is a close relation between the concept of approximate reasoning (Section 7.2), the fuzzy inference engines (Section 8), the fuzzy mapping introduced in the previous section and a fuzzy graph \( \tilde{F} \). Fuzzy rules and a fuzzy graph may both be interpreted as granular representations of functional dependencies and relations.

A fuzzy graph \( \tilde{F} \), serves as an approximate or compressed representation of a functional dependence \( f : X \rightarrow Y \), in the form

\[
\tilde{F} = A_1 \times B_1 \lor A_2 \times B_2 \lor \cdots \lor A_{n_R} \times B_{n_R}
\]

or more compactly

\[
\tilde{F} = \bigvee_{i=1}^{n_R} A_i \times B_i ,
\]

where the \( A_i \) and \( B_i \) are fuzzy subsets of \( X \) and \( Y \), respectively, \( A_i \times B_i \) is the Cartesian product of \( A_i \) and \( B_i \), and \( \lor \) is the operation of disjunction which is usually taken to be the union. In terms of membership functions, we may write

\[
\mu_{\tilde{F}}(x, y) = \bigvee_{i} \left( \mu_{A_i}(x) \land \mu_{B_i}(y) \right) ,
\]

where \( x \in X, y \in Y \), \( \lor \) and \( \land \) are any triangular \( t \)- and \( t \)-conorm, respectively. Usually \( \lor = \max \) and \( \land = \min \) establishing the relationship to the extension principle, approximate reasoning and so forth. A fuzzy graph may therefore be represented as a fuzzy relation or a collection of fuzzy if-then rules

\[
\text{IF } x \text{ is } A_i, \text{ THEN } y \text{ is } B_i \quad i = 1, 2, \ldots, n_R .
\]

Each fuzzy if-then rule is interpreted as the joint constraint on \( x \) and \( y \) defined by

\[
(x, y) \text{ is } A_i \times B_i .
\]

In Section 4.6 we referred to such constraint as a fuzzy point in the data space \( \Xi = X \times Y \). The concept of a fuzzy points and a fuzzy graph are illustrated in Figure 11.3.
11.4 FUZZY LOGIC

Any useful logic must concern itself with Ideas with a fringe of vagueness and a Truth that is a matter of degree.

—Norbert Wiener

The basic claim of fuzzy theorists is that everything is a matter of degree. This view is based on the analysis of 'real-world' problems and the resulting paradox when applying conventional mathematics rooted in set theory. The following examples were first discussed by Bertrand Russell:

**Liar paradox:** Does the liar from Crete lie when he says that all Cretans are liars? If he lies, he tells the truth. But if he tells the truth, he lies.

**Barber paradox:** A barber advertises: “I shave all, and only, those men who don’t shave themselves”. Who shaves the barber? If he shaves himself, then according to the ad he doesn’t. If he does not, then according to him he does.

**Set as a collection of objects:**

i. Consider the collection of books. This collection itself is not a book, thus it is not a member of another collection of books.

ii. The collection of all things that are not books is itself not a book and therefore a member of itself.

iii. Now consider the set of all sets that are not members of themselves. Is this a member of itself or not?
The problem of those paradoxes is their self-reference; they violate the laws of non-contradiction or the law of the excluded middle

\[ t(S) \wedge t(\neg S) = 0 \quad \text{or} \quad S \cap S^c = \emptyset , \]

where \( t \) denotes the truth value (in \( \{0, 1\} \)), \( S \) is a statement or its set representation, respectively. In those paradoxes we find

\[ t(S) = t(\neg S) . \quad (11.13) \]

With \( t(\neg S) = 1 - t(S) \) inserted into (11.13), we obtain the contradiction

\[ t(S) = 1 - t(S) . \quad (11.14) \]

However, in fuzzy logic we simply solve (11.14) for \( t(S) \):

\[ 2 \cdot t(S) = 1 \quad \text{or} \quad t(S) = \frac{1}{2} . \]

...the truth lies somewhere in between!

Throughout the book, we have suggested that fuzzy sets and fuzzy rule-based systems are a suitable means to capture and process vague and fuzzy information. In Section 5, the fuzzy partition \( A_{i1}, \ldots, A_{ir} \) in the if-antecedent part of a rule was identified by projecting fuzzy clusters onto subspaces. In Section 6, the identified model provided forecasts as fuzzy restrictions and we suggested that such information can be readily processed in a fuzzy rule-based system as detailed in Section 7.2. It should, however, be noted that we have not considered fuzzy logic as a multi-valued generalization of a pure (formal) logic. Instead, we motivated our rule-based systems in terms of approximate reasoning (see Section 7.2). Next, we provide more examples of how a fuzzy partition and fuzzy data can be generated from numerical data sets.

### 11.5 A BIJECTIVE PROBABILITY - POSSIBILITY TRANSFORMATION

Possibility theory is an information theory which is related to but independent of both fuzzy sets and probability theory. Technically, a possibility distribution is a fuzzy set. For example, all fuzzy numbers are possibility distributions. However, possibility theory can also be derived without reference to fuzzy sets. Many of the rules of possibility theory are similar to probability theory, but use either max/min or max/times calculus, rather than the plus/times calculus of probability theory. Research into possibility theory has been largely concerned with being precise about uncertainty and reasoning in the presence of uncertainty. Although not as well defined as probability theory, possibility theory has made important contributions to establish formal relationships between a number of paradigms including belief functions, rough sets, random
sets, fuzzy sets and probabilistic sets. The publications of Didier Dubois provide a rich and comprehensive source for most aspects of possibility theory. Here we note only one particular aspect, a formal link between probability and possibility distribution.

A possibility distribution \( \pi(\cdot) \) on \( X \) is a mapping from the reference set or universe \( X \) into the unit-interval,

\[
\pi : X \rightarrow [0,1] .
\]

A usual convention is to assume that there exists at least one \( x \in X \) for which \( \pi(x) = 1 \). This is called the normalization condition. Like a probability distribution is related to its associated probability measure, the possibility distribution is described in terms of a possibility measure by

\[
\pi(x) = \Pi(\{x\}) ,
\]

where the possibility of some event \( A \) is defined by

\[
\Pi(A) = \sup_{x \in X} \pi(x) \quad \text{for an ordinary set } A ,
\]

\[
= \sup_{x \in X} \min\{\mu_A(x), \pi(x)\} \quad \text{for a fuzzy set } A .
\]

(2.14)

A dual necessity (or certainty) measure is defined by

\[
Ne(A) = 1 - \Pi(A^c) ,
\]

\[
= \inf_{x \in A^c} \{1 - \pi(x)\} \quad \text{for an ordinary set } A ,
\]

\[
= \inf_{x \in X} \min\{\mu_A(x), (1 - \pi(x))\} \quad \text{for a fuzzy set } A .
\]

(11.15)

(11.16)

(11.17)

In analogy to axioms in probability theory, \( Pr(X) = \Pi(X) = 1, Pr(\emptyset) = \Pi(\emptyset) = 0 \). The probability of the statement "A or B" is given by \( Pr(A \cup B) = Pr(A) + Pr(B) \). The events \( A \) and \( B \) are assumed to be mutually exclusive, that is, \( A \cap B = \emptyset \). If \( A^c \) denotes the event "not A", in probability theory it is consequently required that probabilities are additive on disjoint sets, \( Pr(A) + Pr(A^c) = 1 \). In contrast, for a subjective judgement one may view both propositions "A occurs" and "A does not occur" as possible and we therefore only require

\[
\Pi(A) + \Pi(A^c) \geq 1 .
\]

(11.18)

To be able to operate with such weak evidence we require the measure of necessity to act as a lower bound:

\[
Ne(A) = 1 - \Pi(A^c) , \quad \text{and} \quad Ne(A) + Ne(A^c) \leq 1 ,
\]

(11.19)

as the impossibility of the opposite event. "A is necessary" is interpreted as \( A \) is bound to occur. We can summarize these intuitive requirements by the following inequality:

\[
Ne(A) = 1 - \Pi(A^c) \leq Pr(A) \leq \Pi(A) .
\]

(11.20)
And for the union of two events,

$$\Pi(A \cup B) = \sup \{\pi(x) : x \in A \cup B\}$$
$$= \sup \{\pi(x) : x \in A \vee x \in B\}$$
$$= \max(\sup\{\pi(x) : x \in A\}, \sup\{\pi(x) : x \in B\})$$

\[ \therefore \Pi(A \cup B) = \max(\Pi(A), \Pi(B)) \] (11.21)

From these definitions, D. Dubois and H. Prade [DP83] derived a necessity measure as a bias in probabilities: the necessity of an event is the extra amount of probability of elementary events in a set over the amount of probability assigned to the most frequent outcome outside the event of concern. For this view, the $x_j$'s are (without loss of generality) ranked so that for $p_j = Pr\{x_j\}$, $p_1 \geq p_2 \geq \cdots \geq p_{n_b}$ and $A_j$ denotes the set $\{x_1, x_2, \ldots, x_j\}$. Thus,

$$Ne(A) = \sum_{x_i \in A} \max(p_i - p', 0), \quad p' = \max_{x_k \notin A} p_k.$$ (11.22)

From (11.22), $\forall j = 1, \ldots, n_b$,

$$\pi_j = 1 - Ne(X - \{x_j\})$$
$$= 1 - \sum_{x_i \in X - \{x_j\}} \max\{p_i - p_j, 0\},$$
$$- p_j + \sum_{i \neq j} (p_i - \max\{p_i - p_j, 0\})$$
$$= p_j - \sum_{i \neq j} \min(p_i, p_j)$$
$$= \sum_{i=1}^{n_b} \min(p_i, p_j).$$ (11.23)

11.6 EXAMPLE: MAINTENANCE DECISION MAKING

One of the major problems in maintenance practice is the lack of a systematic, adaptable and data-driven approach in setting preventive maintenance instructions. A computerized maintenance management system (CMMS) is a decision support system prioritizing machines based on the two most important criteria - the downtime of machines and the frequency of faults. The knowledge base consisting of nine fuzzy rules is shown in Figure 11.4. For example, if a machine fails regularly, but the downtime is relatively small, that is, it does not take much time to solve the problem, the operator should obtain a skill upgrade to deal with the problem himself rather than calling out a maintenance engineer. On the other hand, a machine which, when it fails, does take long to repair, should be monitored for such faults if they do
Fig. 11.4 Decision grid - fuzzy rules used in preventive maintenance decision making.

not occur too frequently. A successful ‘control’ of the maintenance problem is achieved if machines move, month by month from the lower right corner of the decision grid to the upper left corner. The problem is to define what is regarded as a “low downtime”? To obtain a fuzzy partition of the two spaces ‘downtime’ and ‘frequency’, the decision support system reviews monthly data of the worst ten machines. For example, consider the monthly evaluation of ten machines (out of 130) representing about 89% of the problems:

<table>
<thead>
<tr>
<th>d</th>
<th>30</th>
<th>20</th>
<th>20</th>
<th>17</th>
<th>16</th>
<th>12</th>
<th>7</th>
<th>6</th>
<th>6</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>f</td>
<td>27</td>
<td>16</td>
<td>12</td>
<td>9</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

Since the downtime of machines presents the most important criteria, as it can readily be related to profit, loss or gain, we shall focus on downtime hereafter. The first step in obtaining a fuzzy partition is to plot a probability distribution obtained from the histogram. The average downtime, here 13.8, for that particular month is used to split the reference space into two halves. A downtime considerably larger than this value should be regarded as “high”. It is also the point for which $\mu_{\text{med}}(d) = 1$. By definition, probability distributions summarize whole sets of machines; in preventive maintenance, however, decisions are made for individual machines. That is, we do not wish to consider whether or not we can (on average) expect a failure of any machine, but on the basis of the data collected, we wish to establish the degree of feasibility or possibility that a particular machine should be considered as having a “high downtime”. We consequently require a transformation of the probabilistic information to a possibility distribution (or fuzzy restriction, which numerically is the same).
Using the bijective transformation (11.23), \( n_b \) refers to the number of bins of the histogram and \( p \) is the relative frequency estimate of the probability for that section of the space \( D \), of downtime. The data, histogram, and the possibility distribution obtained from (11.23) are shown in Figure 11.5. Finally, fuzzy partitions with piecewise-linear membership functions is obtained from a least-squares fit through the midpoints of the bars of the possibility distribution \( \pi \). For values \( d \) smaller than the mean value, the least-squares algorithm is constrained by the requirement that \( \mu_{med}(d = 0) = 0 \). If we require the sets of the fuzzy partition to be fully overlapping, that is, for any \( d \), \( \sum_i \mu_i(d) = 1 \), the membership functions \( \mu_{low}(d) \) and \( \mu_{high}(d) \) are obtained from the complement of the fuzzy set "medium", \( \mu_{high}(d) = 1 - \mu_{med}(d) \) for values of \( d \) greater than the mean value and \( \mu_{low}(d) = 1 - \mu_{med}(d) \) for values of \( d \) smaller than the mean value.

Fig. 11.5 Data, histogram, possibility distribution and fuzzy partition.

A (multi-stage fuzzy) control or optimization problem is introduced by considering time. Machines are reviewed on a monthly basis and, as indicated by the arrow in Figure 11.4. We aim to move (control) machines from the bottom-right to the upper-left corner of the decision grid.

11.7 EXAMPLE: EVALUATING STUDENT PERFORMANCES

In this example, we wish to evaluate student performances against set standards. At many universities, two thresholds are of particular importance. The student must obtain an overall result of, say \( \geq 50\% \) to obtain the degree. Students with \( \geq 70\% \) are awarded a degree with distinction. It is the neighborhood of these particular points we consider. Strict adherence to such hard thresholds is inevitably unfair in some cases where someone with, say 69.4%,
may not achieve a distinction while a colleague with just 0.6% more would succeed. Let \((X, d)\) be a metric space with \(X = [0, 100]\) and the standard metric
\[
d: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}^+ \\
(x, x') \mapsto |x - x'|.
\]
From (7.6), we know that the metric \(d(\cdot)\) induces a similarity relation \(\tilde{E}: X \times X \rightarrow [0, 1]\) on \(X\) by
\[
\tilde{E} = 1 - \inf(d(x, x'), 1), \tag{7.6}
\]
which formally is identical to a fuzzy restriction or fuzzy set. To identify ‘border cases’, that is, students which are close to the pass mark, denoted \(x_0\), we introduce a function that depends on \(x_0\) and a ‘locality’ parameter \(\delta \in (0, \infty)\) used to scale the metric \(d(\cdot)\) to become a proximity measure. As a result of these considerations, we obtain the triangular fuzzy set
\[
\mu_{x_0}(x) = 1 - \min\{|\delta \cdot x_0 - \delta \cdot x|, 1\}. \tag{11.24}
\]
Figure 11.6 (left plot), shows a fuzzy set induced by \(\delta = 0.1\). A similar procedure can be applied to the 70% mark. Assuming fully overlapping sets \(\forall x \sum \mu(x) = 1\), the single parameter \(\delta\) induces a complete fuzzy partition. Figure 11.6 (right plot) illustrates a general example of a fuzzy partition in performance evaluation.

Let us now consider the results of 23 students in a particular exam. Their results in % are:

<table>
<thead>
<tr>
<th>47.33</th>
<th>34.5</th>
<th>71</th>
<th>84.17</th>
<th>76.5</th>
<th>39.67</th>
<th>62</th>
<th>55.83</th>
</tr>
</thead>
<tbody>
<tr>
<td>35.67</td>
<td>81.67</td>
<td>65</td>
<td>78.83</td>
<td>60.33</td>
<td>43</td>
<td>40.5</td>
<td>40.67</td>
</tr>
<tr>
<td>37.33</td>
<td>64.33</td>
<td>62.67</td>
<td>50</td>
<td>56.67</td>
<td>46.83</td>
<td>26.67</td>
<td></td>
</tr>
</tbody>
</table>

Analyzing the data, we wish to get a picture of how the class performed as a group compared to other classes and how results are distributed within a
group. The overall average of marks is 55%. To get an idea of how results are distributed within the class, we plot a histogram, with bars centered around the points 25, 35, 45, 55, 65, 75, 85. The frequency count is obtained as the number of results that lie in bins from 20 to 90 in steps of 10. The histogram, plotted on the left in Figure 11.7, suggests that the class is somewhat split into two halves. With only few data, one has to be careful as the division of the interval [0, 100] to plot the histogram influences the shape considerably. For instance, the 50% result is counted toward the (40,50] interval. Now, relative frequencies are frequently used to estimate probabilities. From the data, can we determine what the probability was to score more than, say 50%? We would calculate such probability from the density function $p(x)$ as shown in the plot on the right in Figure 11.7 (obtained from Parzen's Gaussian kernel estimator with window width 5; cf. Section 3.2). Considering the definition of statistical laws, the meaning of such evaluation would be senseless, and the plot of the density function does not reveal any useful information.

![Figure 11.7](image)

**Fig. 11.7** Examination results: Histogram and density estimate.