Studies in Fuzziness and Soft Computing

Enric Trillas
Luka Eciolaza

Fuzzy Logic

An Introductory Course for Engineering Students

Springer

Studies in Fuzziness and Soft Computing

Volume 320

Series editor

Janusz Kacprzyk, Polish Academy of Sciences, Warsaw, Poland
e-mail: kacprzyk@ibspan.waw.pl

About this Series

The series "Studies in Fuzziness and Soft Computing" contains publications on various topics in the area of soft computing, which include fuzzy sets, rough sets, neural networks, evolutionary computation, probabilistic and evidential reasoning, multi-valued logic, and related fields. The publications within "Studies in Fuzziness and Soft Computing" are primarily monographs and edited volumes. They cover significant recent developments in the field, both of a foundational and applicable character. An important feature of the series is its short publication time and world-wide distribution. This permits a rapid and broad dissemination of research results.

More information about this series at http://www.springer.com/series/2941

Enric Trillas · Luka Eciolaza

Fuzzy Logic

An Introductory Course for Engineering Students

Springer

Enric Trillas
European Centre for Soft Computing
Mieres, Asturias
Spain

Luka Eciolaza
European Centre for Soft Computing
Mieres, Asturias
Spain

ISSN 1434-9922 ISSN 1860-0808 (electronic)
Studies in Fuzziness and Soft Computing
ISBN 978-3-319-14202-9 ISBN 978-3-319-14203-6 (eBook)
DOI 10.1007/978-3-319-14203-6

Library of Congress Control Number: 2014958577

Springer Cham Heidelberg New York Dordrecht London

Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media (www.springer.com)

This book is dedicated to Professor Lotfi A. Zadeh, as a humble homage of the authors in the 50 years of his paper 'Fuzzy Sets'!

Preface

This book was thought as a non-conventional first course textbook in Fuzzy Logic for engineers ending with an introduction to one of the most fruitful topics arisen from it, Fuzzy Control. It is from the teaching's strategy of the authors, summarized by "Nothing can substitute the own homework of the student" from which it comes its non-conventional character, partially manifested by the 'continuous' form of presenting the considered topics by joining theoretical explanations and examples, and not always following the typically mathematical style of 'theorem-corolaries'.

Behind this strategy is the opinion that, at the university level, students and professors ought to learn jointly, students do not wait to receive everything from the professor's lectures, but should read more than a single recommended textbook. Consequently, this book is neither a manual with recipes to be uncritically applied, nor it is directed to those that can be only interested in mathematical subtleties. The reader should be aware that fuzzy logic is the study and computational management of imprecision and non-random uncertainty, both with the highest accuracy and precision possible at each case, that fuzzy logic is not fuzzy in itself.

Each university course requires a particular teaching tactic that not only depends on the number of lecturing hours, but on the aim of the course and on the audience's characteristics. In particular, additional tutorials supplied by the professor are essential for a good learning process. Tutorials in which other forms of considering the course's topics and more sophisticated problems can be proposed. This is at the own hands of the professor.

The book just presents some basic mathematical models for fuzzy logic but without the intention to just subordinate it to mathematics. Fuzzy logic is neither a part of mathematics, nor even of logic, like Physics is not so. Notwithstanding, what is paramount is the importance and usefulness of mathematical models in experimental sciences and technology, as well as in computer science and computer technology and, in particular, in Soft Computing, where fuzzy logic plays a pivotal role. But the suitability of such models only can come from the success of its testing against some reality, for instance, in true applications; applications play in the techno-scientific world an analogous role to that of experimentation in natural sciences. For instance, if the branch called 'Fuzzy Control' served not as a direct justification of fuzzy logic, the success fuzzy logic has in control applications can be seen as a kind of experimentation to show its usefulness in the study of dynamical systems linguistically described by systems of imprecise rules. Fuzzy logic is much more than what is in this introductory textbook; its applications spread along many domains of science and technology.

The reader should be early acquainted with the fact that, differently from classical bi-valuate logic, almost all in fuzzy logic are context-dependent and purpose-driven. For instance, when representing a system of imprecise linguistic rules in fuzzy terms, all the predicates, connectives, and conditionals in the rules should be specified accordingly with its contextual meaning and the type of inference, forwards or backwards, to be done. In fuzzy logic not only everything is a matter of degree, but its practice requires the art of designing systems by means of the available theoretic armamentarium. The student should be conscious that a mistake in the design process can conduct to solve a different problem than the targeted one.

It should never be forgotten that it is most important in research to not stop questioning (Albert Einstein). Posing good questions (Isaac Rabi) whose answers can result in fertile ones (Karl Menger) is what reveals the relevance of a researcher.

Contents

Chapter 1
On the Roots of Fuzzy Sets

As fuzzy sets were introduced by Zadeh in 1965, they were born closely linked with imprecise predicates, that is, with names of non-precisely defined classes of objects. Even more, most of the applications of Zadeh's ideas are made with properties the objects do verify in some degree between the two classical extremes 0 and 1, respectively. Because of that, it is not at all odd to introduce fuzzy sets from some considerations on how predicates are used in language. We will follow Wittgenstein's statement "The meaning of a word is its use in language".

1.1 A Genesis of Fuzzy Sets

Usually, isolated words mean nothing. To mean something, words are to be used in a given context and in a known way. Words do serve to describe perceptions, to translate reasoning, and to show the reasons for judgements.

For example, what it is meant by the predicate *unleaty*? It is impossible to answer this question, since nobody has never heard something like "this is unleaty", "such is leaty", etc. Neither unleaty, nor leaty, are English words, nobody has used them and they don't appear in English dictionaries. Meaning is inherited by predicates P only after being used, in some ground, by means of elemental statements 'x is P'. By now, *unleaty* has no meaning.

Words are introduced in a language by using them in concrete ways. For example, the Spanish word 'madre' comes from the latin 'mater' that, at its turn, came from one in older indo-european languages, but always used to name someone's mother. Later on, the word could take, by analogy, new meanings as it is, for example, 'mother country', 'goodmother', 'mother in law', 'mother of all wars', 'mother Nature', 'mother-of-pearl', etc. The predicate *leaty* neither appears in English, nor in Spanish, French, German, …, because it was never before used to name a property of the elements in some class, like it is with *tall, young, middle-aged*, with people, *high* with buildings or mountains, or *heavy* with metals, for example. It is only through its use that predicates do acquire 'meaning'.

E. Trillas and L. Eciolaza, *Fuzzy Logic*, Studies in Fuzziness and Soft Computing 320, DOI 10.1007/978-3-319-14203-6_1

What is it meant by the predicate odd? In principle, it depends on where it is used. With natural numbers n, 'n is *odd*' is used accordingly with the mathematical definition/rule 'n is odd if and only if once divided by 2 the rest is 1'. With people, things, or situations, odd does coincide with the meaning of 'strange', 'separated', and 'not often'. In the first context, the predicate odd is precise or crisp, since natural numbers are only odd or not odd, but in other contexts the predicate is imprecise or fuzzy, since for example, 'this is an odd book', or 'that is an odd event', could admit degrees depending on to what extent the book or the event can be qualified as odd. When the predicate is imprecise there is not a perfect classification of the objects to which it refers to.

Only after a predicate acquires meaning, concepts like 'tallness', 'oddity', 'heaviness', etc., appear in the corresponding language. Predicates do appear in language after being used, and only after being used they can evolve in new contexts and give birth to concepts. Like it happened with 'high' and 'highness', and 'royal highness'. Concepts like uncertainty come from a mother-predicate as it is 'uncertain' in this case.

Notice that there are not natural numbers that can be qualified as 'very odd', or 'more or less odd', but there are buildings that can be called to be so. If the predicate is crisp, that is, it names just an either yes, or not, property, the application to it of a linguistic modifier needs to be newly defined, but if the predicate is imprecise it does not, since people immediately understand what it is meant by, for instance, 'very odd'.

In what follows we will only deal with predicates P, the name of a property on a previously given set $X = \{x, y, z, \ldots\}$, and considering the use of P through the elemental statements 'x is P', for all x in X, and accepting (à la Wittgenstein) that the meaning of P is its use in the current language. Hence, the first problem to tackle is placed by the following question, *How the use*, or meaning, *of P on X can be mathematically represented or modeled*?

First, and to distinguish between two statements 'x is P' and 'y is P', for $x \neq y$ in X, let us suppose it is possible to decide when is 'x is less P than y', where x shows the property named P less than y shows it.

Let us call $\leqslant_P$ the relation in X given by

$$x \leqslant_P y \Leftrightarrow x \text{ is less } P \text{ than } y,$$

and suppose $\leqslant_P$ is a preorder (enjoys the reflexive and transitive properties). That is,

- $x \leqslant_P x$, for all x in X
- If $x \leqslant_P y$, and $y \leqslant_P z$, then $x \leqslant_P z$.

The preordered set $(X, \leqslant_P)$ reflects the organization P induces in X, and *the preorder $\leqslant_P$ is the primary use of P in X*. The relation $\leqslant_P^{-1}$ is defined by '$x \leqslant_P^{-1} y \Leftrightarrow y \leqslant_P x$'. Note that the relation $\leqslant_P$ is, usually, empirically perceived; it could also be called the perceptive meaning of P in X.

We will say that 'x is equally P than y' whenever $x \leqslant_P y$ and $y \leqslant_P x$, and write $x =_P y$, with $(=_P) = (\leqslant_P \cap \leqslant_P^{-1})$. Since, obviously,

- $x =_P x$, for all x in X
- $x =_P y \Leftrightarrow y =_P x$
- $x =_P y$, and $y =_P z$ imply $x =_P z$,

the relation $=_P$ is an equivalence in X, and gives the quotient set $X/ =_P$, of the classes of equally-P elements. The predicate P is semi-rigid in X if $X/ =_P$ consists in a finite number of classes. Of course, all predicates on a finite X are semi-rigid.

Example 1.1.1 Let it be $X = \{x_1, \ldots, x_5\}$, and P a predicate inducing the preorder given by the matrix with entries

$$entry\ (i, j) = \begin{cases} 1, & \text{if} \quad x_i \leqslant_P x_j \\ 0, & \text{otherwise,} \end{cases}$$

that is,

$$[\leqslant_P] = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \end{pmatrix}$$

The quotient set $X/ =_P$ has the two classes $\{x_1, x_3, x_5\}$ and $\{x_2, x_4\}$.

If, in the same X, it is Q with primary use defined by

$$x_i \leqslant_Q x_j \Leftrightarrow i \leqslant j,$$

it results $X/ =_Q$ with the 5 classes $\{x_1\}, \{x_2\}, \{x_3\}, \{x_4\}, \{x_5\}$.

Example 1.1.2 In $X = [0, 10]$, consider the predicate $P =$ *around five*, with $\leqslant_P$ defined by

- If $x, y \in [0, 5]$, $x \leqslant_P y \Leftrightarrow x \leqslant y$
- If $x, y \in (5, 10]$, $x \leqslant_P y \Leftrightarrow y \leqslant x$
- If $x \in [0, 5]$, $y \in (5, 10]$, x and y are not $\leqslant_P$-comparable.

Obviously, $\leqslant_P$ is a preorder, and

$$x =_P y \Leftrightarrow \left\{ \begin{array}{l} (x \leqslant y)\&(y \leqslant x) : x = y \\ (y \leqslant x)\&(x \leqslant y) : y = x \end{array} \right\} \Leftrightarrow x = y,$$

that is, $=_P$ is the equality. Hence, $X/ =_P = \{\{x\}; x \in [0, 10]\}$, and P is not semi-rigid.

1.1.1 L-Degree

Let us suppose P in X given by its primary use $\leqslant_P$, and let $(L, \leqslant)$ be a poset. An L-degree, or L-measure, for P in X is a function $\mu_P : X \to L$, such that if $x \leqslant_P y$, then $\mu_P(x) \leqslant \mu_P(y)$- in the order of the poset. The idea behind this definition is that $\mu_P(x) \in L$ evaluates up to which extent x is P, up to which extent x verifies the property named by P. It can be written,

$$\text{Degree up to which } x \text{ is } P = \mu_P(x) \in L,$$

and said that the primary use of P is gradable in $(L, \leqslant)$. If $x =_P y$, then $\mu_P(x) = \mu_P(y)$, as is easily proven. Hence, μ_P is constant in the equivalence's classes modulo $=_P$.

Once μ_P is known, it can be defined the relation $\leqslant_{\mu_P}$ in X by

$$x \leqslant_{\mu_P} y \Leftrightarrow \mu_P(x) \leqslant \mu_P(y),$$

with which it is $\leqslant_P \subset \leqslant_{\mu_P}$. When $\leqslant_P = \leqslant_{\mu_P}$, μ_P *perfectly reflects* the primary use of P in X.

Relation $\leqslant_{\mu_P}$ is a preorder since it is obviously reflexive and transitive. The pair $(\leqslant_P, \leqslant_{\mu_P})$ reflects *a use* of P in X, and once $\leqslant_P$ and $(L, \leqslant)$ are fixed, there can exist again several uses of P in X depending on the L-degree μ_P. L-degrees are also called L-$x5$.

When $L = [0, 1]$, endowed with the linear order of the unit interval, the L-sets are known as fuzzy sets. In this case, the triplet $(X, \leqslant_P, \mu_P)$ is a numerical quantity, where of course, given P and X, neither the relation $\leqslant_P$, nor the measure μ_P are in general unique.

When $L = \{a + ib; 0 \leqslant a \leqslant 1 \ \& \ 0 \leqslant b \leqslant 1\} = \mathbb{C}$ is endowed with the partial order

$$a_1 + ib_1 \leqslant a_2 + ib_2 \Leftrightarrow a_1 \leqslant a_2 \ \& \ b_1 \leqslant b_2,$$

L-sets are 'complex fuzzy sets'. Notice that $(\mathbb{C}, \leqslant)$ is isomorphic with the set of the sub-intervals $[a, b] \subseteq [0, 1]$, once partially ordered by

$$[a_1, b_1] \leqslant [a_2, b_2] \Leftrightarrow a_1 \leqslant a_2 \ \& \ b_1 \leqslant b_2.$$

When it is only known that the degree up to which "x is P" belongs to some sub-interval $[a, b]$, it can be taken $\mu_P(x) = a + ib$. In science and technology, complex quantities are not at all rare.

Remark 1.1.3 Provided $L = [0, 1]$, the relation $\leqslant_{\mu_P}$ is linear or total, since given x, y in X, it is either $\mu_P(x) \leqslant \mu_P(y)$, or $\mu_P(x) \geq \mu_P(y)$; that is, either $x \leqslant_{\mu_P} y$, or $y \leqslant_{\mu_P} x$.

Remark 1.1.4 The relation $\leqslant_P$ is not always linear, there can exist elements x, y in X such that it is neither $x \leqslant_P y$, nor $y \leqslant_P x$, that is, elements which are not comparable under $\leqslant_P$. Of course, in the cases in which $\leqslant_P$ is not linear and $L = [0, 1]$, it cannot be $\leqslant_P = \leqslant_{\mu_P}$, the degree cannot perfectly reflect the primary use of P, and the relation $\leqslant_{\mu_P}$ enlarges the primary use $\leqslant_P$ with all the links between elements in X contained in $\leqslant_P - \leqslant_{\mu_P}$, provided this difference-set is non empty.

Remark 1.1.5 When $(L, \leqslant)$ is not linear, as it is with $L = \mathbb{C}$ the unit complex interval, it is also not sure that $\leqslant_{\mu_P}$ can coincide with the primary meaning, but being not linear there can exist more possibilities for the coincidence.

Remark 1.1.6 Since in fuzzy logic it is always only considered the membership function and not the primary meaning serving for its design, the working scientist should not confuse the two relations $\leqslant_{\mu_P}$ and $\leqslant_P$. For this reason, $\leqslant_{\mu_P}$ can be called the *working-meaning* of P in X, without forgetting that it can be bigger than the primary meaning.

1.1.2 Fuzzy Sets

Once the use $(\leqslant_P, \leqslant_{\mu_P})$ of P in X is given for the poset $(L, \leqslant)$, it will be said that the new object $\underset{\sim}{P}$ defined by

- $x \in_r \underset{\sim}{P}$, if and only if $r = \mu_P(x)$, for $r \in L$,
- $\underset{\sim}{P} = \underset{\sim}{Q}$, if and only if $\mu_P = \mu_Q$

is the L-set labeled P. The set $L^X = \{\mu; \mu : X \to L\}$, is usually and abusively called the set of all L-sets in X, since it contains all the possible degrees in L.

Notice that a predicate P could give many L-sets, in dependance on which poset $(L, \leqslant)$, and which function $\mu_P \in L^X$, are chosen.

From now on, it will be supposed that $(l, \leqslant)$ has a *minimum* element α $(\alpha \leqslant r, \forall r \in L)$, and a *maximum* element $\omega (r \leqslant \omega, \forall r \in L)$. With $L_0 = \{\alpha, \omega\}$, $(L_0, \leqslant)$ is a poset isomorphic to $(\{0, 1\}, \leqslant)$.

Since a classical subset A of X is characterized by its membership function

$$\mu_A(x) = \begin{cases} 1, & \text{if} \quad x \in A \\ 0, & \text{if} \quad x \notin A, \end{cases}$$

the subset A can be viewed as an L-set, whose function only takes the values α or ω, that is,

$$\mu_A(x) = \begin{cases} \omega, & \text{if} \quad x \in A \\ \alpha, & \text{if} \quad x \notin A. \end{cases}$$

Hence, the classical subsets of X are nothing else than the functions in L_0^X, a set included in L^X. Classical subsets are limiting or particular case of L-sets. They are degenerate L-sets.

Notice that if the classical subset A is labeled by a crisp predicate P, this predicate is semi-rigid since $X/=_P$ has, at most, two classes. Crisp predicates are rigid.

Very often μ_P is perceptually designed, that is, designed from what is perceived or known on the concrete use of P in X. It can be thought that such design is purely subjective, in the sense of being made just by what the designer believes on the use of P. But this would not be the case, except if the designer proceeds in a non rational way. The *designer* should try to be as most sure as possible on the correction of his/her/its perceptions.

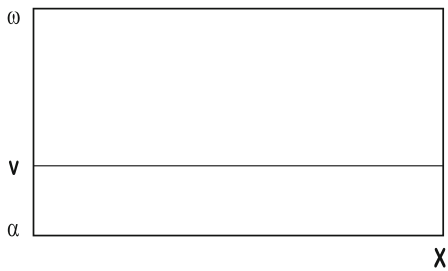
In a lot of cases, mainly in the applications, predicates P in X result indirectly evaluable in $([0, 1], \leqslant)$ thanks to a *numerical characteristic* Ch_P of the current use of P in X. Such characteristic allows to translate 'X is P' into '$Ch_P(x)$ is Q', with an adequate predicate Q on the range of the function $Ch_P : X \to R$. For example, if X is a population and $P = old$, it could be the case that $Ch_{old} =$ *numerical age* $=$ *Age*, in which case 'x is *old*' can be translated into '$Age(x)$ is *big*', with a modeling of *big* according with the current use of *old*, and provided this predicate only depends in the subject's age. In this cases, once the designer is sure that Ch_P and Q are good enough for the case, and also that the order in the numerical interval where Ch_P ranges is adequate to model the order $\leqslant_P$ by the order $\leqslant_Q$, he/she/it should again be sure that the degrees $\mu_Q(Ch_P(x))$ agree with the expected degrees $\mu_P(x)$.

A *design's process* never can arrive to something "exact", but approximate and, if possible, keeping everything under some bounds. For example, by taking the degrees $\mu_Q(Ch_P(x))$ in intervals $(a(x), b(x))$, where $a(x)$ is the minimum of the acceptable values for $\mu_Q(Ch_P(x))$, and $b(x)$ the maximum of them. This can allow to take $\mu_Q(Ch_P(x))$ as, for instance, an average of $a(x)$ and $b(x)$, or as the complex number $a(x) + ib(x)$. For example, if m_x is the middle point of the interval $(a(x), b(x))$, and the confidence that the value is between $a(x)$ and m_x can be quantified in a coefficient a_1, and that of being between m_x and $b(x)$ by a_2, it can be taken the average $\mu_Q(Ch_P(x)) = [a_1 \cdot a(x) + a_2 \cdot b(x)]/a_1 + a_2$.

A rational design should be carefully made by taking into account all the available information, or knowledge, in the use of P in X, as well as of that induced on Q in its numerical universe. If the use of P in X is not known, it is impossible to design neither $\mu_Q(Ch_P)$, nor μ_P, nor to accept $\mu_P(x) = \mu_Q(Ch_P(x))$, for all $x \in X$.

As it was said, if $[x]$ is a class in $X/=_P$, μ_P is constant in it. Let us denote by v_x the value of μ_P in the class $[x]$.

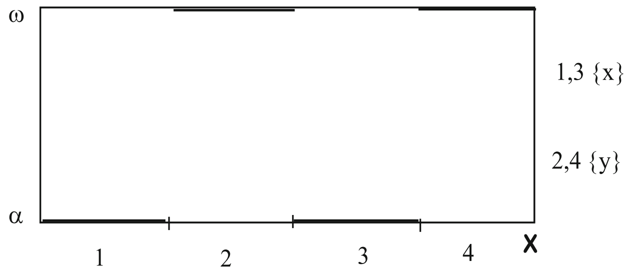
Provided P is semi-rigid with at most two classes in $X/=_P$, then either $X/=_P = \{[x]\}$, or $X/=_P = \{[x], [y]\}$. In the first case, μ_P only has a single value $v \in L$.



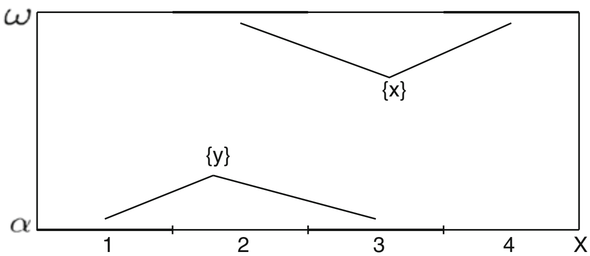
In the second case, μ_P has at most two values, and, for obvious reasons, we will only consider the situation where this values v_x, v_y are different. When,

- $X/=_P = \{[x]\}$, and either $\mu_P(x) = \alpha$ for all x in X, or $\mu_P(x) = \omega$ for all x in X, it results $\underset{\sim}{P} = \varnothing$ in the first case, and $\underset{\sim}{P} = X$ in the second. In both cases, P is a rigid or binary predicate in X.

- $X/=_P = \{[x],[y]\}$, and $v_x \neq v_y$, $\underset{\sim}{P}$ is of the type



In the particular case $\{v_x, v_y\} = \{\alpha, \omega\}$, P is of the type



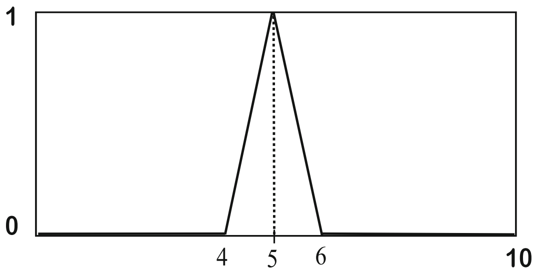
that is, $\underset{\sim}{P}$ is the classical or crisp subset $[x]$ of X, and P is a rigid or binary predicate in X.

Let us show some examples.

Example 1.1.7 With $P =$ *around five*, and $\leqslant_P$ as shown in the example [2 in 1.2], the function

$$\mu_P(x) = \begin{cases} 0, & \text{if} \quad x \in [0,4] \cup [6,10] \\ x-4, & \text{if} \quad x \in [4,5] \\ 6-x & \text{if} \quad x \in [5,6], \end{cases}$$

whose graphics is



verifies: $x \leqslant_P y \Rightarrow \mu_P(x) \leqslant \mu_P(y)$, and hence μ_P is an [0, 1]−degree for P in [0, 10], that originates *a fuzzy set* $\underset{\sim}{P}$ of the numbers in [0, 10] that are around five. Obviously, this degree does not perfectly reflect the primary use of P.

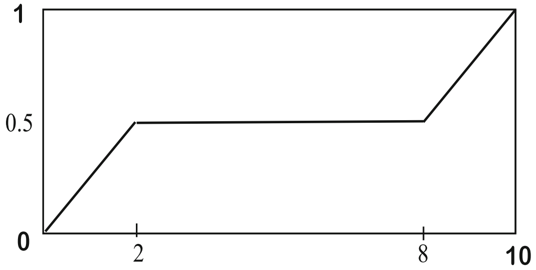
Example 1.1.8 Consider $P = big$ in $X = [0, 10]$. Let us show several possible degrees for it, after agreeing that '$x \leqslant_P y$ if and only if $x \leqslant y$, in the linear order of [0, 10]'.

A [0, 1]-degree μ_P is any function $X = [0, 10] \to [0, 1]$, such that

$$\text{If } x \leqslant y, \text{ then } \mu_P(x) \leqslant \mu_P(y),$$

that is, any non-decreasing function (of which there are many). We can also agree that $\mu_P(0) = 0$, and $\mu_P(10) = 1$. With this, it is clear that all degrees for *big* will show some *family resemblance*.

Once fixed $(L, \leqslant) = ([0, 1], \leqslant)$, and $\leqslant_P = \leqslant$, the different uses of big only depend on which function μ_P is chosen to reflect the meaning of the predicate in [0, 10]. Of course, it is $\leqslant_P = \leqslant_{\mu_P}$ if and only if function μ_P is strictly non-decreasing, as it is the case either with $\mu_P(x) = x/10$, or with $\mu_P(x) = x^2/100$. Provided μ_P is not strictly non-decreasing as, for example, with



since $6 \leqslant_P 4$ (strictly), but $\mu_P(6) = \mu_P(4) = 0.5$, μ_P does not perfectly reflect the primary use of *big* in [0, 10]. In the same way, the crisp degree

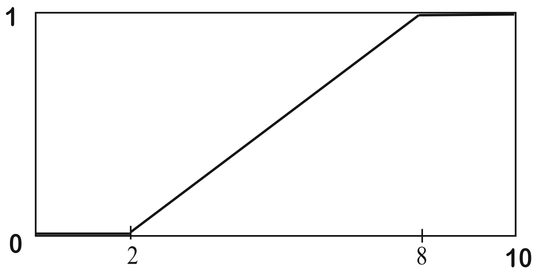
$$\mu_P(x) = \begin{cases} 1, & \text{if} \quad x > 8 \\ 0, & \text{otherwise,} \end{cases}$$

does not perfectly reflect the primary use of *big* when translated into 'after eight'.

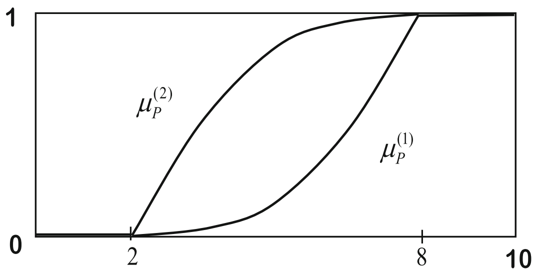
Another possible model for μ_P is

$$\mu_P(x) = \begin{cases} 0, & \text{if} \quad x \in [0, 2] \\ \frac{x-2}{6}, & \text{if} \quad x \in [2, 8] \\ 1 & \text{if} \quad x \in [8, 10], \end{cases}$$

with graphic.



All these models are linear, with the exception os $\mu_P(x) = x^2/100$, that is quadratic. Another quadratic models are given by $\mu_P^{(1)} = (\frac{x-2}{6})^2$, and $\mu_P^{(2)} = 1 - \mu_P^{(1)}(10-x) = 1 - (\frac{8-x}{6})^2$, with graphics

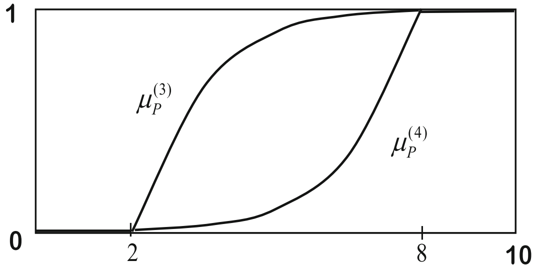


Finally, the following two models are hyperbolic,

$$\mu_P^{(3)}(x) = \begin{cases} 0, & \text{if} \quad x \in [0, 2] \\ \frac{7}{6} \cdot \frac{x-2}{x-1}, & \text{if} \quad x \in [2, 8] \\ 1 & \text{if} \quad x \in [8, 10], \end{cases}$$

$$\mu_P^{(4)}(x) = \begin{cases} 0, & \text{if} \quad x \in [0, 2] \\ -\frac{1}{6}(\frac{7}{x-9} + 1), & \text{if} \quad x \in [2, 8] \\ 1 & \text{if} \quad x \in [8, 10], \end{cases}$$

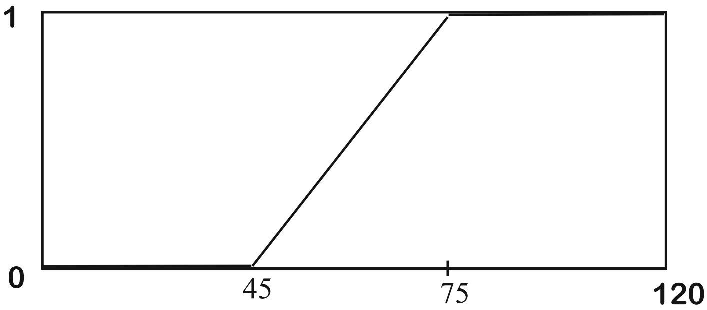
with graphics.



Example 1.1.9 The predicate *old*, once numerically characterized by $Ch_{old} = Age$, can be translated into the interval [0, 120], in years, by $\mu_{old}(x) = \mu_{big}(Age(x))$, and a linear model for it can be

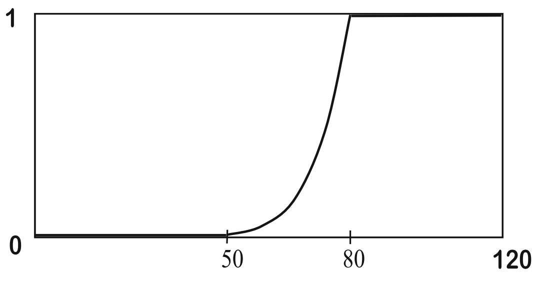
$$\mu_{old}(x) = \mu_{big}(Age(x)) = \begin{cases} 0, & \text{if} \quad 0 \leqslant Age(x) \leqslant 45 \\ \frac{x-45}{30}, & \text{if} \quad 45 \leqslant Age(x) \leqslant 75 \\ 1 & \text{if} \quad 75 \leqslant Age(x) \leqslant 120. \end{cases}$$

with graphic.



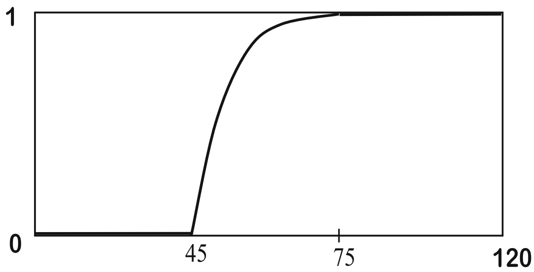
$\mu_{old} = \mu_{big} \circ Age$

In cases like this one, it is important to notice that the function μ_{old} depends on the values 45 and 75, as well as on the form of the curve in the sub-interval [45, 75]. It can be supposed, for example, that the above function is supplied by a person in the range of the fifties, but that one in the seventies would design μ_{old} as the curve



with μ_{old} increasing quadratically between the 50 and the 80 years.

Analogously, a person in the twenties could design μ_{old} as

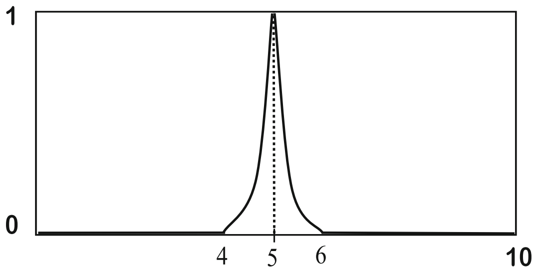


Remark 1.1.10 There are different models for the uses of the same predicate P in X, and such uses are reflected in the corresponding models μ_P in $[0, 1]^X$. It is because of this that it is actually important the process of designing the membership functions.

Example 1.1.11 Analogously to the case of *big*, the predicate $A5 =$ *around five* in $[0, 10]$, can have non-linear but quadratic models, as the one given by

$$\mu_{A5}(x) = \begin{cases} 0, & \text{if} \quad x \in [0, 4] \cup [6, 10] \\ (x-4)^2, & \text{if} \quad x \in [4, 5] \\ (6-x)^2 & \text{if} \quad x \in [5, 6], \end{cases}$$

whose graphics is



Remark 1.1.12 Since each P in a set X can have different degrees μ_P, at each particular case the meaning of P should be well captured to not represent it by a mistaken function that will translate a different use of the predicate.

Remark 1.1.13 Given μ_P, and the $L-$set $\underset{\sim}{P}$, the degree is also called the *membership function* of the L-set. At this respect,

- $x \in_\alpha \underset{\sim}{P}$, is classically written $x \notin \underset{\sim}{P}$
- $x \in_\omega \underset{\sim}{P}$, is classically written $x \in \underset{\sim}{P}$.

1.2 Opposite, Negate, and Middle

Very often the meaning of a predicate P is not captured without simultaneously capturing one of its opposites aP (a for *antonym*, a synonym of *opposite*). How can I recognize that John is *young* without the possibility of recognizing that Peter is *old*? Can someone recognize that a person is *tall* but not that other person is *short*?

The mastering of perception-based predicates shows this kind of polarity: we jointly learn the meaning of P and some of its opposites aP. Even more, without knowing how to use *young* and *old* it is not possible to know how to use *middle-aged*, that is equivalent to '*not young and not old*'. The same could be said with respect to *warm* that is equivalent with '*not cold and not hot*', in relation with water's temperature. Composite predicates of this type are very frequent, for example *medium*, actually equivalent to '*not big and not small*'.

It should be noticed that a 'middle' predicate only exists with imprecise predicates, but not with precise ones. For example, in the set of natural numbers, if $P = even$, it is $aP = odd$, and $a(odd) = even$, thus (*not even*) *and* (*not odd*) $=$ *odd and even*, but for no n it can be stated 'n is *odd and even*'.

Let us remark that, although P and aP are linguistic terms, *not* P is not a linguistic term. For example, in all dictionary we will find *poor* and *rich*, but neither *not poor* nor *not rich*. The negate of P, not P, is more a logical concept than a linguistic one. Our current problem is how to find the uses of $aP(\leqslant_{aP}, \mu_{aP})$, and *not* $P(\leqslant_{not\,P}, \mu_{not\,P})$, given a use $(\leqslant_P, \mu_P)$ of P.

1.2.1 Antonyms

Concerning every opposite aP of P, this opposition is translated by

$$\leqslant_{aP} = \leqslant_P^{-1}$$

since 'x is *less* aP *than* y' should be equivalent to 'y is *less* P *than* x'.

Hence, $\leqslant_{a(aP)} = \leqslant_{aP}^{-1} = (\leqslant_P^{-1})^{-1} = \leqslant_P$, that reflects $a(aP) = P$. For example, with $P = tall$, it is $aP = short$ and $a(aP) = tall$.

This property of aP shows a way for obtaining μ_{aP} once μ_P is known. Let it $A : X \to X$ be a symmetry on X, that is a function such that

- If $x \leqslant_P y$, then $A(y) \leqslant_P A(x)$
- $A \circ A = \mathrm{id}_X$,

and, once $\mu_P : X \to L$ is known, take $\mu_{aP}(x) = \mu_P(A(x))$, for all x in X, that is, $\mu_{aP} = \mu_P \circ A$.

Function $\mu_{aP} = \mu_P \circ A$ is a degree for aP, since:

- $x \leqslant_{aP} y \Leftrightarrow y \leqslant_P x \Rightarrow A(x) \leqslant_P A(y) \Rightarrow \mu_P(A(x)) \leqslant \mu_P(A(y))$, and verifies,

- $\mu_{a(aP)} = \mu_{(aP)} \circ A = (\mu_P \circ A) \circ A = \mu_P \circ (A \circ A) = \mu_P \circ id_X = \mu_P$.

Then, for all symmetry A in X, we have an opposite for P. For example, in $X = [0, 10]$ with $\mu_{big}(x) = \frac{x}{10}$, and $A(x) = 10 - x$, it is $\mu_{big}(10 - x) = \frac{10-x}{10} = 1 - \frac{x}{10}$, and with $a(big) = small$, it can be said $\mu_{small}(x) = 1 - \frac{x}{10}$.

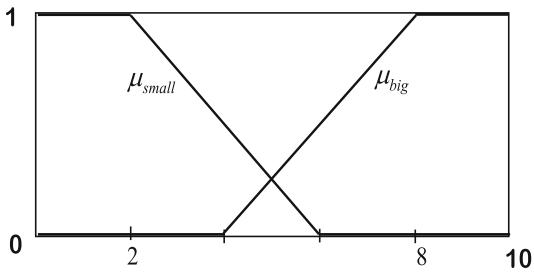
If big is represented by

$$\mu_{big}(x) = \begin{cases} 0, & \text{if} \quad x \in [0, 4] \\ \frac{(x-4)}{4}, & \text{if} \quad x \in [4, 8] \\ 1, & \text{if} \quad x \in [8, 10], \end{cases}$$

with the same symmetry $A(x) = 10 - x$, it results

$$\mu_{small}(x) = \mu_{big}(10 - x) = \begin{cases} 0, & \text{if} \quad x \in [6, 10] \\ \frac{(6-x)}{4}, & \text{if} \quad x \in [2, 6] \\ 1, & \text{if} \quad x \in [0, 2], \end{cases}$$

graphically,



It is easy to prove that $A(x) = 10\frac{10-x}{10+x}$ is also a symmetry in [0, 10], since it verifies,

- If $x \leqslant y$, then $A(y) \leqslant A(x)$,
- $A(A(x)) = x$, for all x in [0, 10]

Hence, another opposite of *big* with $\mu_{big}(x) = x$, is $\mu_{a\,big}(x) = \mu_{big}(A(x)) = A(x) = 10\frac{10-x}{10+x}$. It gives a different representation for *small*, $\mu_{small}(x) = 10\frac{10-x}{10+x}$ in [0, 1].

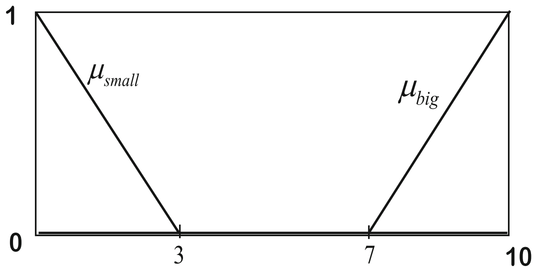
The last two examples show a serious trouble. There are no points $x \in X$ such that it is simultaneously $\mu_P(x) = 0$ and $\mu_{aP}(x) = 0$. The pairs of opposites (P, aP) for which there is a region in X such that both μ_P and μ_{aP} take the value 0, called neutral region, are called *regular opposites*. Hence, the two above pairs (big, small) are *not regular*. Nevertheless, if *big* is represented by

$$\mu_{big}(x) = \begin{cases} 0, & \text{if} \quad 0 \leqslant x \leqslant 7 \\ \frac{(x-7)}{3}, & \text{if} \quad 7 \leqslant x \leqslant 10, \end{cases}$$

with $A(x) = 10 - x$, results

$$\mu_{small}(x) = \mu_{big}(10 - x) = \begin{cases} 0, & \text{if} \quad 3 \leqslant x \leqslant 10 \\ \frac{(3-x)}{3}, & \text{if} \quad 0 \leqslant x \leqslant 3, \end{cases}$$

with graphics



that shows the neutral region (3, 7). This pair is regular.

1.2.2 Negations

Let it P be a predicate in X, and $P' = notP$ its negate. The only we can say about the relation between $\leqslant_P$ and $\leqslant_{P'}$ is that it is $\leqslant_{P'} \subset \leqslant_P^{-1}$, since

$$\text{If } x \text{ is } \textit{less P than } y, \text{ then } y \text{ is } \textit{less not P than } x,$$

or, equivalently, $\leqslant_P \subset \leqslant_{P'}^{-1}$. We can also easily agree that,

- If $\mu_P(x) = \alpha$, then $\mu_{P'}(x) = \omega$
- If $\mu_P(x) = \omega$, then $\mu_{P'}(x) = \alpha$.

Let it $N : L \to L$ be a function such that

1. If $a \leqslant b$, then $N(b) \leqslant N(a)$,
2. $N(\alpha) = \omega$, and $N(\omega) = \alpha$,

with such a function N, it is $\mu_{P'} = N \circ \mu_P$ an L-degree for P', since

$$x \leqslant_{P'} y \Rightarrow y \leqslant_P x \Rightarrow \mu_P(y) \leqslant \mu_P(x)$$
$$\Rightarrow N(\mu_P(x)) \leqslant N(\mu_P(y)) \Leftrightarrow \mu_{P'}(x) \leqslant \mu_{P'}(y).$$

Hence, given an L-degree μ_P of P in X, with each function N verifying (1) and (2), we get the L-degree $\mu_{P'} = N \circ \mu_P$. Such functions N are called *negation functions*.

Provided the negation function does verify

3. $N \circ N = \mathrm{id}_L$,

then

$$\mu_{(P')'}(x) = N(\mu_{P'}(x)) = N(N(\mu_P(x)))$$
$$= (N \circ N)(\mu_P(x)) = \mathrm{id}_L(\mu_P(x)) = \mu_P(x),$$

for all x in X, or $\mu_{(P')'} = \mu_P$.

Functions N verifying (1), (2), and (3) are called *strong negations*, and are almost the only used with fuzzy sets, for the following reason. With $L = [0, 1]$, all strong negations are, obviously, continuous, hence, if μ_P is continuous also $\mu_{P'}$ is such, and if μ_P has some discontinuities, $\mu_{P'}$ has the same discontinuities. That is, strong negations do not add discontinuities to those of μ_P.

If $L = [0, 1]$, there is a family of strong negations widely used in fuzzy set theory, the so-called Sugeno's negations:

$$N_\lambda(a) = \frac{1-a}{1+\lambda a}, \text{ with } \lambda > -1, \quad \text{for all} \quad a \in [0, 1].$$

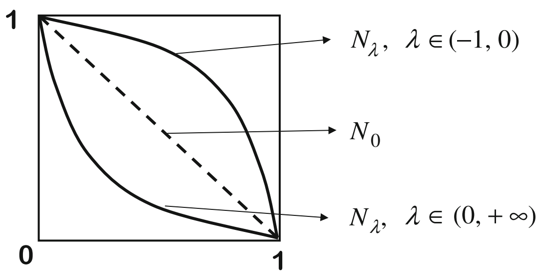
For example, $N_0(a) = 1 - a$, $N_1(a) = \frac{1-a}{1+a}$, $N_{-0.5} = \frac{1-a}{1-0.5a}$, $N_2(a) = \frac{1-a}{1+2a}$, etc. Since obviously,

$$N_{\lambda_1} \leqslant N_{\lambda_2} \Leftrightarrow \lambda_1 \leqslant \lambda_2,$$

it results:

- If $\lambda \in (-1, 0]$, then $N_\lambda \leqslant N_0$
- If $\lambda \in (0, +\infty]$, then $N_0 < N_\lambda$,

graphically



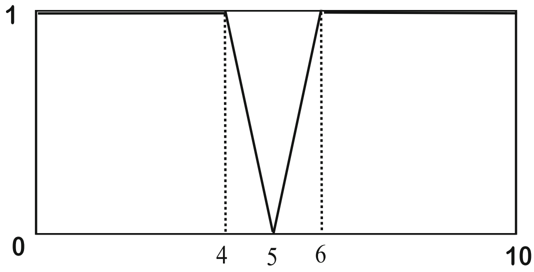
Notice that, provided N is in the Sugeno's family of strong negations, it is enough to know a concrete pair of numbers $(a, N_\lambda(a))$ to compute the corresponding λ. For example,

- If $N_\lambda(0.5) = 0.5$, it results $\lambda = 0$
- If $N_\lambda(0.7) = 0.4$, it results $\lambda = -\frac{5}{14}$
- If $N_\lambda(0.4) = 0.5$, it results $\lambda = \frac{1}{2}$.

Example 1.2.1 With μ_{A5} as defined in 1.1.6, and with N_0, it is

$$\mu_{not\ A5}(x) = 1 - \mu_{A5}(x) \begin{cases} 1, & \text{if} \quad x \in [0, 4] \cup [6, 10] \\ 5 - x, & \text{if} \quad x \in [4, 5] \\ x - 5 & \text{if} \quad x \in [5, 6], \end{cases}$$

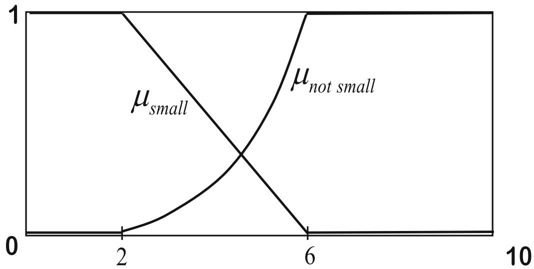
with graphics.



Example 1.2.2 With μ_{small} as defined in the second case of 2.2, and with N_1 it is

$$\mu_{not\ small}(x) = N_1(\mu_{small}(x)) = \frac{1 - \mu_{small}(x)}{1 + \mu_{small}(x)} = \begin{cases} 1, & \text{if} \quad 6 \leqslant x \leqslant 10 \\ \frac{x-2}{10-x}, & \text{if} \quad 2 \leqslant x \leqslant 6 \\ 0 & \text{if} \quad 0 \leqslant x \leqslant 2, \end{cases}$$

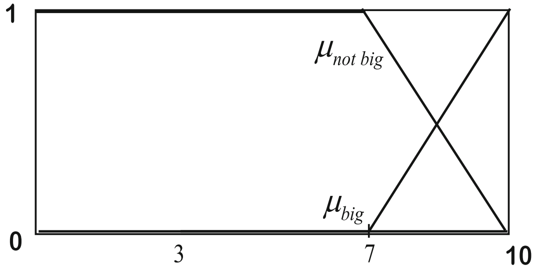
whose graphics are



Example 1.2.3 With μ_{big} as defined in the third case of 2.2, and with N_0, it is

$$\mu_{not\ big}(x) = \begin{cases} 1, & \text{if} \quad 0 \leqslant x \leqslant 7 \\ \frac{10-x}{3}, & \text{if} \quad 7 \leqslant x \leqslant 10, \end{cases}$$

with graphics



1.2.3 Antonyms and Negations

With pairs of antonyms (P, aP), it should always be taken into account that conditional statements like "If the bottle is empty, then it is not full", conduct to the inequality $\mu_{aP} \leqslant \mu_{not\ P}$, with $P = full$, showing that *not* P *could be taken as the biggest antonym of* P. It is not often the case in which $aP = not\ \ P$, practically it only happens when aP is such that there is not any linguistic term aP in the language. When aP and *not* P are not coincidental, it is said that aP is a *strict antonym* of P.

When modeling $\mu_{aP} = \mu_P \circ A$, and $\mu_{not\ P} = N \circ \mu_P$, with a symmetry A of X and a strong negation N in [0, 1], it results the *condition of coherence*:

$$\mu_P \circ A \leqslant N \circ \mu_P,$$

between A and N, that should be always verified. If A is known, N should be chosen to satisfy this coherence's condition, and if what is known is N, then A should be chosen to verify such condition.

Example 1.2.4 With $N_0(a) = 1 - a$, and $A(x) = 10 - x$ in $X = [0, 10]$, if

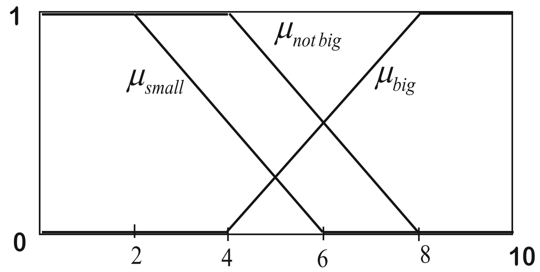
$$\mu_{big}(x) = \begin{cases} 0, & \text{if} \quad x \in [0, 4] \\ \frac{x-4}{4}, & \text{if} \quad \in [4, 8] \\ 1, & \text{if} \quad x \in [8, 10], \end{cases}$$

results

$$\mu_{small}(x) = \mu_{big}(10 - x) = \begin{cases} 0, & \text{if} \quad x \in [6, 10] \\ \frac{6-x}{4}, & \text{if} \quad \in [2, 6] \\ 1, & \text{if} \quad x \in [0, 2], \end{cases}$$

$$\mu_{not\ big}(x) = 1 - \mu_{big}(x) = \begin{cases} 1, & \text{if} \quad x \in [0, 4] \\ \frac{8-x}{4}, & \text{if} \quad \in [4, 8] \\ 0, & \text{if} \quad x \in [8, 10], \end{cases}$$

whose graphics show that the pair (*big, small*) is coherent, since $\mu_{small} \leqslant \mu_{not\ big}$.

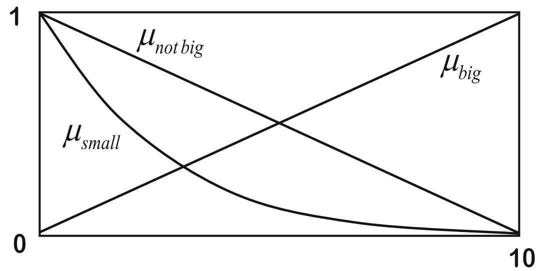


Example 1.2.5 In [0, 10] take $\mu_{big}(x) = \frac{x}{10}$ and $\mu_{not\ big}(x) = 1 - \mu_{big}(x) = 1 - \frac{x}{10}$. Which symmetries $A : [0, 10] \to [0, 10]$ can be taken for having $\mu_{small} = \mu_{big} \circ A$? From the coherence's condition with N_0, follows

$$\mu_{small}(x) = \mu_{big}(A(x)) = \frac{A(x)}{10} \leqslant \mu_{not\ big}(x) = 1 - \frac{x}{10}$$

hence, $A(x) \leqslant 10 - x$ is the condition A must satisfy. For example,

- If $A_1(x) = 10 - x$, it results $\mu_{small}(x) = 1 - \frac{x}{10} = \mu_{not\ big}(x)$, a non-regular case.
- If $A_2(x) = 10 \cdot \frac{10-x}{10+x}$, for which $A_2(x) \leqslant 10 - x$, it results $\mu_{small}(x) = \mu_{big}(10 \cdot \frac{10-x}{10+x}) = \frac{10-x}{10+x}$, with graphics



also showing coherence.

Example 1.2.6 With the same $\mu_{big}(x) = \frac{x}{10}$ in the previous example, and $\mu_{small} = \mu_{big}(10-x)$, which Sugeno's strong negation N_λ can be used for having $\mu_{not\ big}(x) = N_\lambda(\mu_{big}(x))$?

It should be,

$$\mu_{small}(x) = \mu_{big}(10 - x) = \frac{10 - x}{10} \leqslant N_\lambda(\mu_{big}(x)) = N_\lambda(\frac{x}{10}) = \frac{10 - x}{10 + \lambda x}$$

that means $\lambda x \leqslant 0$. Hence, $\lambda \in (-1, 0]$. For example,

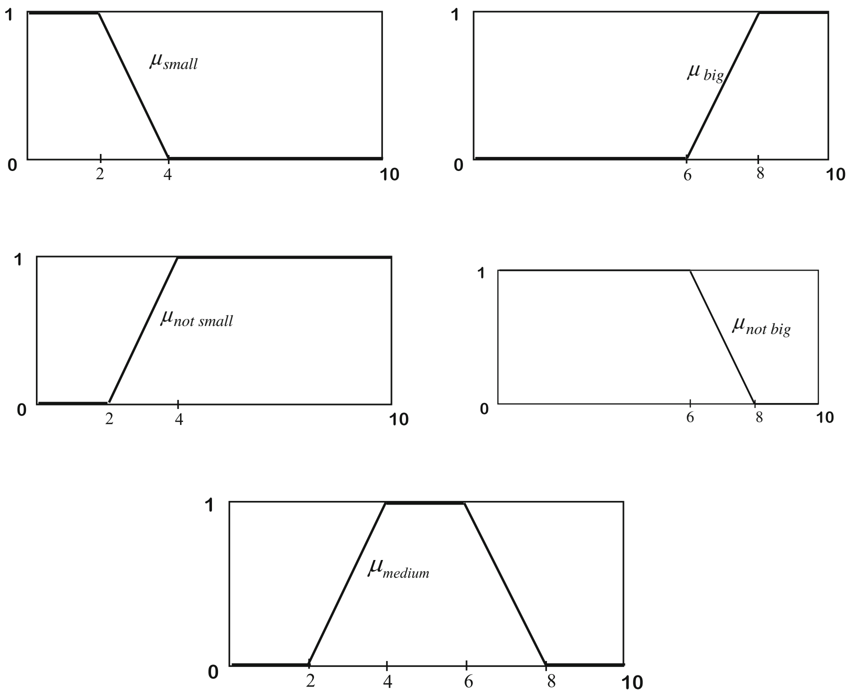
- With $N_\lambda = N_0$, there is coherence
- With $N_\lambda = N_{-0.5}$, or $N_\lambda(a) = \frac{2(1-a)}{2-a}$, there is coherence
- With $N_\lambda = N_1$, there is no coherence.

1.2.4 Medium Term

Given a regular pair of antonyms (P, aP), the middle or medium term of them, is the predicate *MP* = *not P and Not aP*, as it was said in Sect. 1.2.1. Although we don't yet studied the diverse models for the conjunction 'and', by the moment let us take the model given by the operation minimum (= min) in [0, 1]. Then,

$$\mu_{MP} = \min(\mu_{not\ P}(x), \mu_{not\ aP}(x)).$$

For example, with $P = small$ in [0, 10], N_0, and $A(x) = 10 - x$, is:



The triplets (P, MP, aP) play an important role in the applications of fuzzy sets.

1.3 AND/OR

Let P, Q be two predicates in X. Consider the new predicates 'P and Q', and 'P or Q', used by means of:

- 'x is P and Q' $\Leftrightarrow$ 'x is P' *and* 'x is Q'
- 'x is P or Q' $\Leftrightarrow$ Not (Not 'x is P' *and* Not 'x is Q') $\Leftrightarrow$ Not ('x is P'' *and* ' x is Q'') $\Leftrightarrow$ x is '(P' *and* Q')',

with respective primary uses $\leqslant_{P\ and\ Q} \subset \leqslant_P \cap \leqslant_Q$, and $\leqslant_{P\ or\ Q}$. Take L-degrees μ_P, μ_Q.

1.3.1 AND

Given $(L, \leqslant)$, let $* : L \times L \to L$ be an operation, verifying the properties

- $a \leqslant b, c \leqslant d \Rightarrow a * c \leqslant b * d$
- $a * c \leqslant a, a * c \leqslant c,$

Then, $\mu_P(x) * \mu_Q(x)$ is an L-degree for P and Q in X, since:
$x \leqslant_{P\ and\ Q} y \Rightarrow x \leqslant_P y$ *and* $x \leqslant_Q y \Rightarrow \mu_P(x) \leqslant \mu_P(y)$ *and* $\mu_Q(x) \leqslant \mu_Q(y) \Rightarrow \mu_P(x) * \mu_Q(x) \leqslant \mu_P(y) * \mu_Q(y)$.

Hence, defining $\mu_P(x) * \mu_Q(x) = \mu_{P\ and\ Q}(x)$, an L-degree for '$P$ and Q' in X is obtained, and the operation $*$ can be called an *and-operation* or a conjunction. Notice that it is,

$$\mu_{P\ and\ Q}(x) \leqslant \mu_P(x), \text{ and } \mu_{P\ and\ Q}(x) \leqslant \mu_Q(x).$$

In the case $(L, \leqslant)$ is a lattice (see Sect. 1.6) with the minimum operation $\cdot = \min$, and $* \leqslant \cdot$, it implies

$$\mu_{P\ and\ Q}(x) \leqslant \mu_P(x) \cdot \mu_Q(x).$$

Then, if $(L, \cdot, +, \leqslant)$ is a lattice, then the *and*-operation is at least $* = \cdot(\min)$, with which we have the L-degree

$$\mu_{P\ and\ Q}(x) = \mu_P(x) \cdot \mu_Q(x), \forall x \in X.$$

1.3.2 OR

If $*$ is an and-operation, and $N : L \to L$ is a strong negation, define

$$a \oplus b = N(N(a) * N(b)), \text{ for all } a, b \in L.$$

Since $a \leqslant b, c \leqslant d \Rightarrow N(b) \leqslant N(a), N(d) \leqslant N(c) \Rightarrow N(b) * N(d) \leqslant N(a) * N(c) \Rightarrow N(N(a) * N(c)) \leqslant N(N(b) * N(d))$, it results $a \oplus c \leqslant b \oplus d$.

Analogously, from $N(a) * N(b) \leqslant N(a)$, it follows $a \leqslant N(N(a) * N(b)) = a \oplus b$, and $b \leqslant a \oplus b$, for all a, b.

Then, $\mu_P(x) \oplus \mu_Q(x)$ is an L-degree for P or Q, since (remember that it is $\leqslant_{P'} \subset \leqslant_P^{-1}$):

$x \leqslant_{P \text{ or } Q} y \Leftrightarrow x \leqslant_{(P' \text{ and } Q')'} y \Leftrightarrow y \leqslant_{P' \text{ and } Q'} x \Rightarrow \mu_{P' \text{ and } Q'}(y) \leqslant \mu_{P' \text{ and } Q'}(x) \Leftrightarrow \mu_{P'}(y) * \mu_{Q'}(y) \leqslant \mu_{P'}(x) * \mu_{Q'}(x) \Rightarrow N(\mu_P(y)) * N(\mu_Q(y)) \leqslant N(\mu_P(x)) * N(\mu_Q(x)) \Rightarrow N(N(\mu_P(x)) * N(\mu_Q(x)) \leqslant N(N(\mu_P(y)) * N(\mu_Q(y)) \Leftrightarrow \mu_P(x) \oplus \mu_Q(x) \leqslant \mu_P(y) \oplus \mu_Q(y)$.

Hence, $\mu_P(x) \oplus \mu_Q(x)$ can be taken as an L-degree for '*P* or *Q*' in *X*, and the operation $\oplus$ can be called an *or-operation* or a disjunction. Notice that

$$\mu_P(x) \leqslant \mu_{P' \text{ or } Q'}(x), \quad \mu_Q(x) \leqslant \mu_{P' \text{ or } Q'}(x).$$

In the case $(L, \cdot, +)$ is a lattice with the maximum operation $+ =$ max, and $+ \leqslant *$, it implies $\mu_P(x) + \mu_Q(x) \leqslant \mu_{P' \text{ or } Q'}(x)$.

Then, if $(L, \cdot, +, \leqslant)$ is a lattice, at least it is the operation $\oplus = +$(max), with which we have the L-degree

$$\mu_{P \text{ or } Q}(x) = \mu_P(x) + \mu_Q(x), \forall x \in X.$$

With such degree it holds the 'duality' law

$$\mu_{P' \text{ or } Q'} = \mu'_{P \text{ and } Q}.$$

Remark 1.3.1 The lattice operation $\cdot(+)$ is not necessarily the only operation $*(\oplus)$, that can exist. In the case $(L, \leqslant)$ is not a lattice for $\cdot(+)$, there can also exist other operations $*$ and $\oplus$. For example, $L = [0, 1]$ is not a lattice with $* = prod$, but

$$a \leqslant b, c \leqslant d \Rightarrow prod(a, c) \leqslant prod(b, d),$$

and $prod(a, b) \leqslant a$, $prod(a, b) \leqslant b$. Hence, $prod$ can be eventually used to model the use of *and*. Analogously, $\oplus = prod^*(a, b) = 1 - prod(1 - a, 1 - b) = a + b - a \cdot b$ verifies $a \leqslant b, c \leqslant d \Rightarrow prod^*(a, c) \leqslant prod^*(b, d)$, and $a \leqslant prod^*(a, b), b \leqslant prod^*(a, b)$. Hence, $prod^*$ can be eventually used to model the use of *or*.

Remark 1.3.2 The existence of operations $*$ and $\oplus$ in L, warrants the existence of L-degrees for *and*, *or*, respectively.

Remark 1.3.3 Since $a * b \leqslant (a * b) \oplus (a * b)$, and $a * b \leqslant a, a * b \leqslant b$ it follows

$$(a * b) \oplus (a * b) \leqslant a \oplus b,$$

and $a * b \leqslant a \oplus b$, for all a, b in L, and all pair of operations $*$ and $\oplus$.

Remark 1.3.4 It should be noticed that what has been presented is sufficient but not necessary for the representation of $\mu_{P\ and\ Q}$ and $\mu_{P\ or\ Q}$ b means of μ_P and μ_Q.

Given P, if with M for *medium*, $MP = Not\ P\ \ and\ \ Not\ aP = P'\ \ and\ \ (aP)'$, with a negation function N for *not*, and a symmetry A for the opposite, and $*$ for *and*, results

$$\mu_{MP}(x) = \mu_{P'\ and\ (aP)'}(x) = \mu_{P'}(x) * \mu_{(aP)'}(x) = N(\mu_P(x)) * N(\mu_P(A(x))),$$

for all x in X.

1.4 Qualified, Modified, and Constrained Predicates

1.4.1 Qualified Predicates

Let P be a predicate on X, with L-degree $\mu_P : X \to L$, and τ a predicate on $\mu_P(X) \subset L$, with L-degree $\mu_\tau : \mu_P(X) \to L$. Suppose $\leqslant\ \subset\ \leqslant_\tau$, and consider the *qualified predicate* 'P is τ',

$$\text{'}x \text{ is } (P \text{ is } \tau)\text{'} := x \text{ is } P \text{ is } \tau,$$

provided $\varnothing \neq\ \leqslant_{P\ is\ \tau}\ \subset\ \leqslant_P$. On these conditions,

$$\mu_{P\ is\ \tau} = \mu_\tau \circ \mu_P$$

is an L-degree for 'P is τ' in X, since:

$$\begin{aligned} x \leqslant_{P\ is\ \tau} y &\Rightarrow x \leqslant_P y \Rightarrow \mu_P(x) \leqslant \mu_P(y) \Rightarrow \mu_P(x) \leqslant_\tau \mu_P(y) \\ &\Rightarrow \mu_\tau(\mu_P(x)) \leqslant \mu_\tau(\mu_P(y)), \end{aligned}$$

that is, $(\mu_\tau \circ \mu_P)(x) \leqslant (\mu_\tau \circ \mu_P)(y)$.

For example, with $L = [0, 1]$, $P = small$ in $[0, 10]$, with $\leqslant_P\ =\ \leqslant^{-1}$, and $\mu_P(x) = 1 - \frac{x}{10}$, if $\tau = large$ in $[0, 1]$ is with $\leqslant_\tau\ =\ \leqslant$, and

$$\mu_\tau(x) = \begin{cases} 0, & \text{if} \quad 0 \leqslant x \leqslant 0.5 \\ 1, & \text{if} \quad 0.5 \leqslant x \leqslant 1, \end{cases}$$

it results

$$\mu_\tau \circ \mu_P(x) = \begin{cases} 0, & \text{if} \quad 5 \leqslant x < 10 \\ 1, & \text{if} \quad 0 \leqslant x < 5, \end{cases}$$

that allow the interpretation of *small is large* as *less than five*.

Take $\tau = true$ on [0, 1], with $\leqslant_\tau = \leqslant$, and the degree μ_τ as a non-decreasing function $[0, 1] \to [0, 1]$, such that $\mu_\tau(0) = 0$, $\mu_\tau(1) = 1$, once accepting that '0 is τ' is false, and '1 is τ' is true. Taking $\mu_\tau(x) = x$, as it is usual in fuzzy logic, it is

$$\text{Degree up to which } x \text{ is } P \text{ is true} = \mu_\tau(\mu_P(x)) = \mu_P(x), \text{ for all } x \text{ in } X,$$

that allows to accept, as it is usual,

$$\text{Degree of true of } x \text{ is } P = \mu_P(x).$$

1.4.2 Linguistic Modifiers

Linguistic modifiers or linguistic hedges, m, are adverbs acting on P just in the concatenated form mP. For example, with $m = very$ and $P = tall$, it is $mP = very\ tall$.

A characteristic that linguistically distinguishes imprecise predicates from precise ones, is that in the first case and once P and m are given, *mP* is immediately understood. If P is precise (for example, $P = even$ in the set of natural numbers), *mP* needs of a new definition to be understood (what it means *very even*?). Modifiers only modify, but do not change abruptly imprecise predicates.

If P in X is with the use $(\leqslant_P, \mu_P)$, and m in $\mu_P(x) \subset L$ is with $\leqslant \subset \leqslant_m$ and μ_m, provided $\leqslant_{mP} \subset \leqslant_P$, it can be taken the degree

$$\mu_{mP} = \mu_m \circ \mu_P,$$

since, $x \leqslant_{mP} y \Rightarrow x \leqslant_P y \Rightarrow \mu_P(x) \leqslant \mu_P(y) \Rightarrow \mu_P(x) \leqslant_m \mu_P(y) \Rightarrow \mu_m(\mu_P(x)) \leqslant \mu_m(\mu_P(y))$, or $(\mu_m \circ \mu_P)(x) \leqslant (\mu_m \circ \mu_P)(y)$.

Among linguistic modifiers there are two specially interesting types:

- *Expansive modifiers*, verifying $\mathrm{id}_{\mu_P(x)} \leqslant \mu_m$,
- *Contractive modifiers*, verifying $\mu_m \leqslant \mathrm{id}_{\mu_P(x)}$.

With the expansive, it results

$$\mathrm{id}_{\mu_P(x)}(\mu_P(x)) = \mu_P(x) \leqslant \mu_m(\mu_P(x)) = \mu_{mP}(x) : \mu_P(x) \leqslant \mu_{mP}(x), \quad \text{for all } x \text{ in } X.$$

With the contractive, it results

$$\mu_{mP}(x) = \mu_m(\mu_P(x)) \leqslant \mathrm{id}_{\mu_P(x)}(\mu_P(x)) = \mu_P(x) : \mu_{mP}(x) \leqslant \mu_P(x), \quad \text{for all } x \text{ in } X.$$

This is what happens in $L = [0, 1]$ with the Zadeh's old definitions,

$$\mu_{more\ or\ less}(a) = \sqrt{a}, \quad \mu_{very}(a) = a^2.$$

1.4.3 Constrained Predicates

Let P and Q predicates in X and Y, respectively, with uses $(\leqslant_P, \mu_P)$, $(\leqslant_Q, \mu_Q)$. Each relation $\emptyset \neq R(P, Q)$:

$$(x \text{ is } P, y \text{ is } Q) \in R(P, Q),$$

allows to define the *constrained predicate* $Q|P =$ 'Q if P', in $X \times Y$, given by

$$(x, y) \in Q|P \Leftrightarrow (x \text{ is } P, y \text{ is } Q) \in R(P, Q).$$

An example is given by the interpretation

$$(x, y) \in Q|P \Leftrightarrow \text{ 'If } x \text{ is } P, \text{ then } y \text{ is } Q' \Leftrightarrow \text{ '}x \text{ is } P \text{ implies } y \text{ is } Q'.$$

Provided $Q|P$ induces a preorder $\leqslant_{Q|P}$ in $X \times Y$, and that there is an L-degree $\mu_{Q|P} : X \times Y \to L$,

$$(x_1, y_1) \leqslant_{Q|P} (x_2, y_2) \Rightarrow \mu_{Q|P}(x_1, y_1) \leqslant \mu_{Q|P}(x_2, y_2),$$

it could be studied how to express $\mu_{Q|P}$ by means of μ_P and μ_Q.

Notice that there are several possibilities for obtaining $\leqslant_{Q|P}$ from both $\leqslant_P$ and $\leqslant_Q$, i.e.,

$$\leqslant_{Q|P} = \leqslant_P \times \leqslant_Q, \leqslant_{Q|P} = \leqslant_P^{-1} \times \leqslant_Q, \text{ etc.}$$

The degree $\mu_{Q|P}$ is said to be decomposable, or functionally expressible, if there is an operation $J : L \times L \to L$, such that

$$\mu_{Q|P}(x, y) = J(\mu_P(x), \mu_Q(y)),$$

for all $(x, y) \in X \times Y$, and it again remains to be tested that $\mu_{Q|P}$ is actually a L-degree for $Q|P$. For example,

- If $\leqslant_{Q|P} = \leqslant_Q \times \leqslant_P$, and J is non-decreasing in both variables, it is $(x_1, y_1) \leqslant_{Q|P} (x_2, y_2) \Leftrightarrow x_1 \leqslant_P x_2, y_1 \leqslant_Q y_2 \Rightarrow \mu_P(x_1) \leqslant \mu_P(x_2)$, and $\mu_Q(y_1) \leqslant \mu_Q(y_2) \Rightarrow$

  $$J(\mu_P(x_1), \mu_Q(y_1)) \leqslant J(\mu_P(x_2), \mu_Q(y_2)),$$

  or

  $$\mu_{Q|P}(x_1, y_1) \leqslant \mu_{Q|P}(x_2, y_2).$$

- If $\leqslant_{Q|P} = \leqslant_P^{-1} \times \leqslant_Q^{-1}$, and J is decreasing in both variables, it also follows the same conclusion,
- If $\leqslant_{Q|P} = \leqslant_P^{-1} \times \leqslant_Q$, and J is decreasing in its first variable, and non-decreasing in the second, it also follows the same conclusion.
  Etc.

Remark 1.4.1 The decomposability or functional expressibility of $\mu_{Q|P}$, $\mu_{not\ P}$, μ_{aP}, $\mu_{P\ and\ Q}$, and $\mu_{P\ or\ Q}$, is not a general property of the L-degrees of the predicates $Q|P$, $not\ P$, aP, $P\ and\ Q$, and $P\ or\ Q$. What has been shown at such respect with functions J, N, A, $*$, and $\oplus$, respectively, is just to be taken as examples of the existence of L-degrees. Although in the applications of fuzzy logic is currently accepted that all these predicates are functionally expressible, that is, expressed through numerical functions

$$J, *, \oplus : [0, 1] \times [0, 1] \to [0, 1], N : [0, 1] \to [0, 1], A : X \to X,$$

it should not be considered that this is always the case.

Remark 1.4.2 $Q|P$ is an example of a *relational predicate*, that is, a predicate R on $X \times Y$ such that $(x, y) \in R$, with $x \in X$, and $y \in Y$. For example, R = larger, implies, around, etc. Of course, once either x or y are fixed, what results is a predicate (unary) in Y or X, respectively, as it is with 'x is around y', if $X = Y = [0, 10]$, where with $y = 5$ it results the unary predicate *around five* in $[0, 10]$.

Relational, or binary, predicates can be either precise or imprecise. In the first case, they originate a crisp subset of $X \times Y$ defined by

$$\mu_R(x, y) = \begin{cases} \omega, & \text{if} \quad (x, y) \in R \\ \alpha, & \text{otherwise.} \end{cases}$$

In the second, they originate an L-set in $X \times Y$ defined by

$$\mu_R(x, y) = \text{ Degree in } L \text{ up to which it is } (x, y) \in R,$$

once an L-degree for R is known.

Remark 1.4.3 In the case $L = [0, 1]$, functions $J : [0, 1] \times [0, 1] \to [0, 1]$ allowing to represent $\mu_{Q|P}$ by $J \circ (\mu_P \times \mu_Q)$, are called *fuzzy relations*, and if the predicate $Q|P$ interprets a rule, these relations are called *fuzzy conditionals*.

1.4.4 Group Meaning

The meaning of words is not fixed for all people and all context. For example, in a dinner with three commensals the deliciousness of the dessert plates could easily result in three different orderings of such plates. Since language is a social phenomenon, also meaning is such, and it is possible to talk on the meaning of predicates for a group of people in, of course, a given context.

For a group of people $G = \{p_1, \ldots, p_m\}$, a predicate P on X can show m primary meanings $\leqslant_{P,i}$, $1 \leqslant i \leqslant m$. Since

$$(\bigcap_{i=1}^{m} \leqslant_{P,i}) = \leqslant_{P,G}$$

is not empty (all $\leqslant_{P,i}$ are reflexive), it can be taken:

$$Primary\ meaning\ of\ P\ on\ X\ for\ the\ group\ G = \leqslant_{P,G}.$$

Notice that provided all $\leqslant_{P,i}$ are preorders, $\leqslant_{P,G}$ is also a preorder.

If $m\ L - degrees\ \mu_P^{(i)}$ are known for each primary meaning $\leqslant_{P,i}$, since

- $x =_{P,G} y \Leftrightarrow x =_{P,1} y \,\&\ldots\&\, x =_{P,m} y$,
- $x \leqslant_{P,G} y \Leftrightarrow x \leqslant_{P,1} y \,\&\ldots\&\, x \leqslant_{P,m} y$,

for each function $\Phi : L^m \to L$, non-decreasing in each place i between 1 and m (for example, if $a \leqslant b$ then $\Phi(a, x_2, \ldots, x_m) \leqslant \Phi(b, x_2, \ldots, x_m)$), or aggregation function, it results

- $x \leqslant_{P,G} y \Rightarrow \Phi(\mu_P^{(1)}(x), \ldots, \mu_P^{(m)}(x)) \leqslant \Phi(\mu_P^{(1)}(y), \ldots, \mu_P^{(m)}(y))$,

that allows to take

$$\mu_P^G(X) = \Phi(\mu_P^{(1)}(x), \ldots, \mu_P^{(m)}(x)), \text{ for all } x \in X,$$

as an *aggregate* $L - degree$ of P on X for the group G. The meaning for G results from aggregating its people's meanings.

1.4.5 Synonims

In the language, synonymy is a complex problem whose roots are possibly to be searched for in the apparition of new facts or concepts for which there is not yet a word for their designation. Then, what is sometimes done is to designate the new fact/concept by means of an old word whose meaning is considered, for some reasons, similar to that of the new fact/concept. That is, for example, that in which the old word was already used in situations judged similar to those where the new fact/concept appears/applies.

Synonymy is related with some kind of similarity or proximity of meaning but here we will only try to present some previous treats of it.

Let P be a predicate on X with $\leqslant_P$, and Q a predicate on Y with $\leqslant_Q$. If there exists a bijective function $u : X \to Y$ such that,

- $x_1 \leqslant_P x_2 \Leftrightarrow u(x_1) \leqslant_Q u(x_2)$,

predicates P and Q are *u-primary-synonyms*. Notice that when $X = Y$, with $u = id_X$, what results is that P and Q are id_X-primary synonyms, or *primary synonyms* for short, if and only if $\leqslant_P = \leqslant_Q$, that is, if and only if

$$\text{Primary meaning of } P \text{ on } X = \text{Primary meaning of } Q \text{ on } X$$

If P and Q are id_X-primary synonyms, it is said that they are exact or perfect synonyms when $\mu_P = \mu_Q$, and it results $(\leqslant_P, \leqslant_{\mu_P}) = (\leqslant_Q, \leqslant_{\mu_Q})$.

For example, if $P = small$ on $X = [0, 1]$ is with $\leqslant_P = \leqslant^{-1}$ (the reverse linear order on the real line), and $Q = short$ on $Y = [0, 10]$ is with $\leqslant_Q = \leqslant^{-1}$ (also the reverse linear order on the real line), $u(x) = 10x$ gives

- $x \leqslant^{-1} y \Leftrightarrow 10x \leqslant^{-1} 10y$

taking $\leqslant_P \cap \leqslant_P^{-1}$ and $\leqslant_Q \cap \leqslant_Q^{-1}$ equal to the identity (=) on the real line. Then, *small* and *short* can be considered a pair of u-primary synonyms.

If P acts on X with an $\mathcal{L} - degree\ \mu_P$, Q acts on Y and is a u-primary synonym of P, from

$$y_1 \leqslant_Q y_2 \Leftrightarrow u^{-1}(y_1) \leqslant_P u^{-1}(y_2) \Rightarrow \mu_P(u^{-1}(y_1)) \leqslant \mu_P(u^{-1}(y_2)),$$

it follows that

$$\mu_Q = \mu_P \circ u^{-1}$$

is an $\mathcal{L} - degree$ for Q. In this situation it is

$$y_1 \leqslant_{\mu_Q} y_2 \Leftrightarrow u^{-1}(y_1) \leqslant_{\mu_P} u^{-1}(y_2),$$

or

$$x_1 \leqslant_{\mu_P} x_2 \Leftrightarrow u(x_1) \leqslant_{\mu_Q} u(x_2),$$

that are equivalent to

$$\leqslant_{\mu_Q} = \leqslant_{\mu_P} \circ (u \times u).$$

For example, with the before mentioned predicates *short* and *small*, it is

$$\mu_{short}(y) = \mu_{small}(y/10)$$

for all y in [0, 10], and results

$$y_1 \leqslant_{\mu_Q} y_2 \Leftrightarrow y_1/10 \leqslant_{\mu_P} y_2/10.$$

Remark 1.4.4 Whenever P and Q are u-synonyms, it could be stated that "*P means Q*".

Remark 1.4.5 The definition of *primary meaning* is just a formal one trying to approach an important aspect of the meaning of linguistic predicates, when acting on a given universe of discourse. The same can be said about the definition of u-primary synonyms with which it does not hold, in general, that a pair of linguistic synonyms are necessarily $u-$primary synonyms. Anyway, what can be said is that Q is a *migration* of P to the universe Y.

Remark 1.4.6 In some way, the current meaning of a predicate, the form in which it is used today in the plane language, partially inherits its past meanings.

It is the evolution of the use of P in different universes that produces the descriptions of P that are found in dictionaries. Such current meaning, or meanings, always never can be described by a 'if and only if' definition like in the case of $P = odd$ in the set of positive integers, but just by some more or less clear description, as it is with the same word 'odd' when applied either to social situations, or to people. Most of the words appearing in dictionaries are imprecise and, hence, not definable by precise terms. For this kind of reasons, dictionaries contain several short descriptions for most of the words in them.

The current meaning of P is also influenced with the simultaneous use of its synonyms and antonyms, as well as by the more or less different uses of P some groups of people did, for saying nothing of the influences coming from other natural languages. Of course, there are some words that evanish since their universe of application disappears; words are 'born' when applied to concrete universes, and are 'dying' when these universes disappear, as it happened, for instance, with some old words that were use in the wide middle age's world of cavalry. Language is a dynamical system, and their dynamicity implies that it cannot be well studied by only employing the methods that are typical of classical logic. Fuzzy logic could be a mathematical and practical tool for that study in the cases where ambiguity can be left aside.

1.5 Linguistic Variables

Linguistic variables are basic tools in most application of fuzzy sets in the technology's field. They do mainly appear when linguistically describing the behavior of the physical variables of a system. A linguistic variable explicits a concept by (linguistically) granulating some elemental components of it, by showing the perceptually distinguishable shades that are relevant for the corresponding application.

A linguistic variable LV is formed after considering

1. Its principal predicate, P
2. One of the opposites of P, aP
3. Some linguistic modifiers $m_1, \ldots, m_n$,
   and by adding:
4. Its negate (*not P*), or the middle-predicate (*MP*), or *P and Q*, or *not* $m_1 P$, or *P and* $m_2 aP, \ldots$

Then LV, is called *the linguistic variable generated by* P, and reflects the linguistic granulation perceived for the concept. For example,

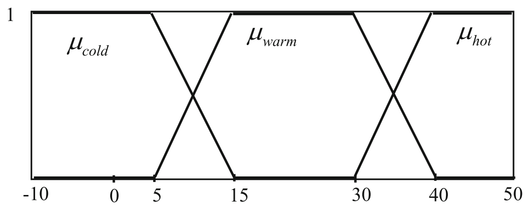
- LV = Age, is Age = { young, old, middle-aged, not old, not very young, ...}
- LV = Truth, is Truth = { true, false, very false, not very true, ...}
- LV = Temperature, is Temperature = { cold, hot, warm, not cold, not very hot, ...}
- LV = Size, is Size = {large, small, medium, very large, ...}
- LV = Height, is

  - For buildings, Height = {low, high, medium, very high, not very low, …}
  - For people, Height = {tall, short, medium, very tall, more or less short, …}

- LV = Speed, is Speed = {fast, slow, very slow, more or less fast, not fast, …}

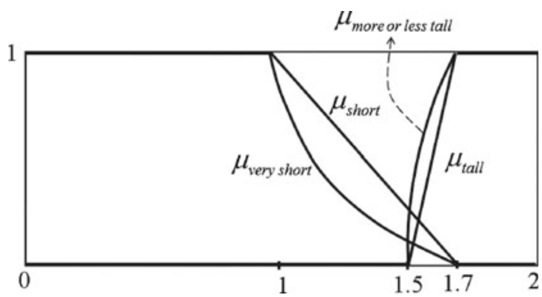
Although the number of terms in a Linguistic Variable can be large, usually it is comprised between 5 and 9 (7 ± 2) since, in a lot of cases, less than 5 shades is poor and more than 9 is excessive. Anyway in a good number of applications there only appear the three terms *P*, *aP*, and *MP*.

Usually, in the applications, the variables range in the set of real numbers and, because of this, the predicate P acts in some interval of the real line. For example, the linguistic variable 'Temperature' in the interval between



-10, and 50 degrees Celsius, is often represented by only the three terms μ_{cold}, μ_{hot}, μ_{warm}, with warm = not cold and not hot.

Analogously, 'Height' for people, can be represented in [0, 2] meters by the linguistic variable with the four terms μ_{tall}, μ_{short}, $\mu_{very\ short}$, $\mu_{more\ or\ less\ tall}$ in the following figure,



1.5.1 Fuzzy Partition

It is sometimes useful in the applications that, once ordered in some sequence, the fuzzy sets in a linguistic variable $LV = \{\mu_0, \mu_1, \ldots, \mu_n\}$ do form what is called a *fuzzy partition* (or a unit's partition), that is, verifying

$$\sum_{j=0}^{n} \mu_j(x) = 1, \forall x \in X.$$

This definition is a direct generalization of what happens in the classical case. If $X = A_0 \cup A_1 \cup \cdots \cup A_n$, with $A_i \cap A_j = \varnothing$ for $i \neq j$, is a classical partition of X, it is $\mu_{A_0}(x) + \mu_{A_1}(x) + \cdots + \mu_{A_n}(x) = 1$, since for each $x \in X$ there is just one A_j such that $x \in A_j$, but $x \notin A_i$, for $i \neq j$, that is,

$$\mu_{A_j}(x) = 1, \text{ and } \mu_{A_i}(x) = 0 \quad \text{if} \quad i \neq j,$$

that implies $\sum_{j=0}^{n} \mu_{A_j}(x) = 1$. Let us show three examples.

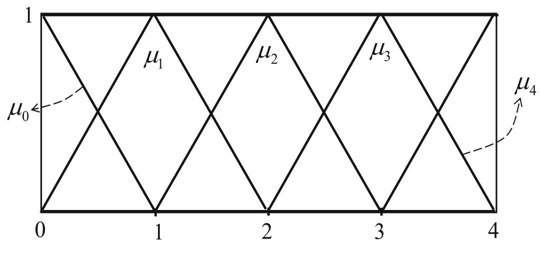
Example 1.5.1 In $X = [0, 4]$, take

$$\mu_0(x) = \begin{cases} 1 - x, & \text{if} \quad 0 \leqslant x \leqslant 1, \\ 0 & \text{if} \quad 1 \leqslant x \leqslant 4, \end{cases}$$

$$\mu_j(x) = \begin{cases} 0, & \text{if} \quad 0 \leqslant x \leqslant j - 1, \\ x + 1 - j & \text{if} \quad j - 1 \leqslant x \leqslant j, \\ j + 1 - x & \text{if} \quad j \leqslant x \leqslant j + 1, \\ 1, & \text{if} \quad j + 1 \leqslant x \leqslant 4, \end{cases} \quad \text{for} \quad 1 \leqslant j \leqslant 3$$

$$\mu_4(x) = \begin{cases} 0, & \text{if} \quad 0 \leqslant x \leqslant 3, \\ x - 3 & \text{if} \quad 3 \leqslant x \leqslant 4, \end{cases}$$

Graphically,



Obviously,

- If $0 \leqslant x \leqslant 1$, $\sum_{j=0}^{4} \mu_j(x) = \mu_0(x) + \mu_1(x) = 1 - x + x = 1$
- If $1 \leqslant j \leqslant 3$, $j \leqslant x \leqslant j + 1$, $\sum_{j=0}^{4} \mu_j(x) = \mu_j(x) + \mu_{j+1}(x) = (j + 1 - x) + (x + 1 - j - 1) = 1$.
- If $3 \leqslant x \leqslant 4$, $\sum_{j=0}^{4} \mu_j(x) = \mu_3(x) + \mu_4(x) = 4 - x + x - 3 = 1$

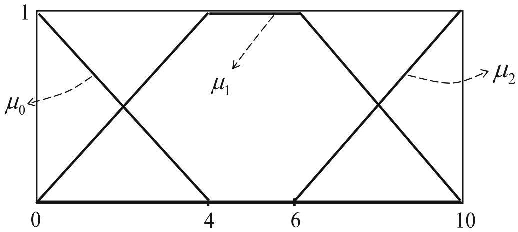
Hence $\{\mu_0, \mu_1, \ldots, \mu_n\}$ is a fuzzy partition of $[0, 4]$. Notice that each μ_j can be labeled by the predicate around j= A_j, that is $\mu_j = \mu_{A_j}$.

Example 1.5.2 In $X = [0, 10]$, take

$$\mu_0(x) = \begin{cases} 1 - \frac{x}{4}, & \text{if } \ 0 \leqslant x \leqslant 4, \\ 0, & \text{if } \ 4 \leqslant x \leqslant 10 \end{cases}, \quad \mu_1(x) = \begin{cases} \frac{x}{4}, & \text{if } \ 0 \leqslant x \leqslant 4, \\ 1, & \text{if } \ 4 \leqslant x \leqslant 6 \\ \frac{10-x}{4}, & \text{if } \ 6 \leqslant x \leqslant 10 \end{cases},$$

$$\mu_2(x) = \begin{cases} 0, & \text{if } \ 0 \leqslant x \leqslant 6, \\ \frac{x-6}{4}, & \text{if } \ 6 \leqslant x \leqslant 10. \end{cases}$$

Graphically



Since, obviously, $\mu_0(x) + \mu_1(x) + \mu_2(x) = 1$ for all $x \in [0, 10]$, $\{\mu_0, \mu_1, \mu_2\}$ is a fuzzy partition of the interval $[0, 10]$.

Example 1.5.3 By proceeding in the same way that in last example, it is easy to prove that it is

$$\{\mu_{cold}, \mu_{hot}, \mu_{warm}\},$$

a fuzzy partition of the interval $[-10, 50]$.

Notice that, with the strong negation $N_0 = 1-\text{id}$, if $\{\mu_0, \mu_1, \mu_2\}$ is a fuzzy partition of X, it is

$$\begin{array}{ll} \mu_1(x) + \mu_2(x) = 1 - \mu_0(x), & \text{or} \quad \mu_1 + \mu_2 = \mu_0' \\ \mu_0(x) + \mu_1(x) = 1 - \mu_2(x), & \text{or} \quad \mu_0 + \mu_1 = \mu_2' \\ \mu_0(x) + \mu_2(x) = 1 - \mu_1(x), & \text{or} \quad \mu_0 + \mu_2 = \mu_1' \end{array}$$

The complement (with N_0) of each element in a fuzzy partition, is just the addition of the other elements in the partition.

1.6 A Note on Lattices

1. An algebraic structure $(L, \leqslant)$ is a partially ordered set, or poset, if the relation $\leqslant \subseteq L \times L$ is a partial order, that is, enjoys the properties:

   - Reflexive, $a \leqslant a$ for all a in L.
   - Transitive, $(a \leqslant b)\ \&\ (b \leqslant c) \Rightarrow (a \leqslant c)$.
   - Anti-symmetric, $(a \leqslant b)\ \&\ (b \leqslant a) \Leftrightarrow a = b$.

Provided $(L, \leqslant)$ only verifies the reflexive and the transitive properties, is called a preordered set.

2. An algebraic structure $(L, \cdot, +)$ is a lattice, provided the binary operations $\cdot$ and $+$ verify,

   - Both operations are commutative and associative, that is, $a \cdot b = b \cdot a, a + b = b + a, a \cdot (b \cdot c) = (a \cdot b) \cdot c, a + (b + c) = (a + b) + c$, for all a, b, c in L.
   - Both operations are idempotent, that is, $a \cdot a = a, a + a = a$, for all a in L.

   Without proof, let's state what follows:

   (a) It is easy to prove that the relation $a \leqslant b \Leftrightarrow a \cdot b = a$, is a partial order called the 'natural order of the lattice', and that it is equivalent to $a \leqslant b \Leftrightarrow a + b = b$.
   (b) Operation $+$ verifies: $a + b$ is the lowest element in $(L, \leqslant)$ such that $a \leqslant a + b$, and $b \leqslant a + b$. That is, in the natural order, there is no element in L that is both between a and $a + b$, and between b and $a + b$. Operation $\cdot$ verifies: $a \cdot b$ is the greatest element in $(L, \leqslant)$ such that $a \cdot b \leqslant a$, and $a \cdot b \leqslant b$. That is, in the natural order, there is no element in L between $a \cdot b$ and a, and between $a \cdot b$ and b, in the order $\leqslant$.
   (c) If $a \leqslant b$, then $a \cdot c \leqslant b \cdot c$, and $a + c \leqslant b + c$, for all c in L.

3. A lattice is bounded if there are elements 0 and 1 in L, such that $0 \leqslant a \leqslant 1$, for all a in L and called, respectively, the minimum and the maximum of the lattice. These elements verify:
   $0 \cdot a = 0$, $1 \cdot a = a$, $0 + a = a$, and $1 + a = 1$, for all a in L. That is, 0 is absorbent for $\cdot$, and neutral for $+$, and 1 is neutral for $\cdot$, and absorbent for $+$.
4. In a bounded lattice, a mapping $' \colon L \to L$ is a negation provided: $0' = 1$, $1' = 0$, $a \leqslant b \Rightarrow b' \leqslant a'$, and $a'' = (a')' = a$, for all a in L.
5. A bounded lattice with a negation is dual, provided the operations $\cdot$ and $+$ are linked by $(a \cdot b)' = a' + b'$, for all a, b in L. This expression is equivalent to $(a + b)' = a' \cdot b'$, since it immediately follows from $(a' \cdot b')' = a'' + b'' = a + b$.
6. A lattice is distributive, provided: $a + b \cdot c = (a + b) \cdot (a + c)$, and $a \cdot (b + c) = a \cdot b + a \cdot c$.
   Provided it is dual, a bounded distributive lattice with a negation is a De Morgan algebra. A bounded lattice with a negation such that it is dual, and: $a \cdot a' = 0$ (law of non-contradiction), and $a + a' = 1$ (law of excluded-middle), is an Ortho-lattice. Notice that both laws are equivalent; for instance, $a \cdot a' = 0 \Rightarrow (a \cdot a')' = 1 \Rightarrow a' + a = 1$.
   An Ortho-lattice such that it holds the property
   $a \leqslant b \Leftrightarrow b = a + a' \cdot b$,
   is called an Ortho-modular lattice. Distributive ortho-lattices, are called Boolean algebras. Notice that in them:

   - $a \leqslant b \Rightarrow a + a' \cdot b = (a + a') \cdot (a + b) = a + b = b$
   - $b = a + a' \cdot b \geq a \Rightarrow a \leqslant b$,

   that is, all Boolean algebras are Ortho-modular lattices, but not reciprocally.

It is obvious that a De Morgan algebra $(L, \cdot, +, ')$ verifying the laws of non-contradiction and excluded-middle is a Boolean algebra. In all De Morgan algebra there is included the Boolean algebra whose elements are the Boolean elements of the algebra, that is, that given by the set $\{a \in L; a \cdot a' = 0\} = \{a \in L; a + a' = 1\}$, a set that is never empty since at least contains 0 and 1.

1.6.1 Examples

1. A good example of a Boolean algebra is the power-set $P(X) = \{A; A \subseteq X\}$, whose elements are the subsets of a set X. In this case, the operations are:
   - the intersection $\bigcap$ of subsets is $\cdot$
   - the union $\bigcup$ of subsets is $+$
   - the complement of subsets $(^c)$, is $'$,

   and 0 is the empty set $\varnothing$, and 1 is the full set X.
   Power-sets are typical instances Boolean algebras. In fact, any Boolean algebra is isomorphic to a power-set.
2. The unit interval [0, 1], endowed with the operations: $\cdot = \min$, $+ = \max$, and $' = 1 - id$, is a De Morgan algebra, but not a Boolean one, since, for instance, $min(a, 1 - a) = 0 \Leftrightarrow a = 0$, or $a = 1$. The Boolean elements of this algebra are just 0 and 1.
3. The set of functions $[0, 1]^X = \{\mu; \mu : X \to [0, 1]\}$, endowed with the operations given by $(\mu \cdot \sigma)(x) = min(\mu(x), \sigma(x))$, $(\mu + \sigma)(x) = max(\mu(x), \sigma(x))$, $\mu'(x) = 1 - \mu(x)$, for all x in X, is a De Morgan algebra, whose Boolean elements are the functions $\mu \in \{0, 1\}^X$. Hence, this algebra is not a Boolean one.
4. The set whose elements are all the vector subspaces of R^3 constitute an Orthomodular lattice, once endowed with the operations:
   - The intersection of two subspaces $(\cdot)$,
   - The minimum subspace that contains two subspaces $(+)$,
   - The subspace that is orthogonal to a subspace $(')$.

   For instance, if it is $< u, v, w > = R^3$, that is, these three vectors are an orthogonal basis of R^3, it is:
   - $< u > \cdot < v > = 0$, the null subspace, with $0 = (0, 0, 0)$,
   - $< u > + < w > = < u, w >$, the plane given by the two orthogonal vectors u an w,
   - $< v >' = < u, w >$, the same plane.

Chapter 2
Algebras of Fuzzy Sets

2.1 Introduction

From now on it will be only considered the case in which $(L, \leqslant) = ([0, 1], \leqslant)$, that is, of Zadeh's fuzzy sets, with predicates P in X known through a degree $\mu_P : X \to [0, 1]$, and without knowing, necessarily, its primary use $\leqslant_P$. The set of all fuzzy sets in X, $[0, 1]^X$, will be also denoted by $F(X)$. In this case, the preorder $\leqslant_{\mu_P}$ is linear, or total, since for all x, y in X it is either $\mu_P(x) \leqslant \mu_P(y)$, or $\mu_P(y) \leqslant \mu_P(x)$, that is, it is either $x \leqslant_{\mu_P} y$ or $y \leqslant_{\mu_P} x$ for all x, y in X. Hence, $\leqslant_{\mu_P}$ rarely will perfectly reflect the primary use of P in X, since $\leqslant_P$ is usually not linear.

In the case in which X is finite, $X = \{x_1, \ldots, x_n\}$, the fuzzy sets $\mu \in [0, 1]^X$, will be represented by

$$\mu = \mu(x_1)/x_1 + \mu(x_2)/x_2 + \cdots + \mu(x_n)/x_n,$$

with the convention that if some term $\mu(x_j)/x_j$ does not appear, is that it is $\mu(x_j) = 0$. For example, with $X = \{1, 2, 3, 4\}$, the expression

$$\mu = 0.5/x_1 + 0.7/x_2 + 1/x_4,$$

refers to the fuzzy set in X given by $\mu(x_1) = 0.5, \mu(x_2) = 0.7, \mu(x_3) = 0, \mu(x_4) = 1$. Analogously, the fuzzy set $\mu' = N_0 \circ \mu$ ($N_0 = 1 - \mathrm{id}$) is

$$\mu' = 0.5/x_1 + 0.3/x_2 + 1/x_3,$$

2.1.1 Cartesian Product

If A, B are crisp subsets in X and Y, respectively, that is, $A \in \mathbb{P}(X)$ and $B \in \mathbb{P}(X)$, its cartesian product $A \times B = \{(a, b); a \in A, b \in B\} \subset X \times Y$, is with the

E. Trillas and L. Eciolaza, *Fuzzy Logic*, Studies in Fuzziness and Soft Computing 320, DOI 10.1007/978-3-319-14203-6_2

membership function $\mu_{A\times B} : X \times Y \to \{0, 1\}$, given by

$$\mu_{A\times B}(x, y) = \min(\mu_A(x), \mu_B(y)$$

for all $x \in X$, $y \in Y$. It is $\mu_{A\times B}(x, y) = 1 \Leftrightarrow \mu_A(x) = \mu_B(y) = 1$.

In the same vein, if $\mu \in F(X)$, and $\sigma \in F(Y)$, the cartesian product $\mu \times \sigma$ is defined by directly generalizing the classical case:

$$\mu \times \sigma = \min \circ (\mu, \sigma).$$

For example, with $X = \{x_1, x_2, x_3\}$, $Y = \{y_1, y_2\}$, and

$$\mu = 1/x_1 + 0.8/x_2,\ \ \sigma = 0.9/y_1 + 0.7/y_2,$$

it is

$$\mu \times \sigma = 0.9/(x_1, y_1) + 0.7/(x_1, y_2) + 0.8/(x_2, y_1) + 0.7/(x_2, y_2),$$

with $(\mu \times \sigma)(x_3, y_1) = (\mu \times \sigma)(x_3, y_2) = 0$.

With $\mu_{big}(x) = \frac{x}{5}$ if $x \in [0, 5]$ and $\mu_{small}(y) = 1 - \frac{y}{7}$ if $y \in [0, 7]$, it is

$$(\mu_{big} \times \mu_{small})(x, y) = \min(\frac{x}{5}, 1 - \frac{y}{7}),$$

the representation of the cartesian product as a surface contained in the cube $[0, 5] \times [0, 7] \times [0, 1]$.

Of course, if $\mu = \mu_A \in \{0, 1\}^X$, $\sigma = \mu_B \in \{0, 1\}^X$, it is not only $\mu \times \sigma \in \{0, 1\}^X$ but $\mu \times \sigma = \mu_{A\times B}$.

2.1.2 Extension Principle

If $f : X \to Y$ is a mapping and A is a crisp subset of X, $A \subset X$, it is $f(A) = \{y \in Y; f(a) = y, a \in A\}$ the f-image of A in Y. Notice that

$$\mu_{f(A)}(y) = \sup\{\mu_A(x); f(x) = y\} = \begin{cases} 1, & \text{if it exists } x \in A \text{ such that } f(x) = y, \\ 0, & \text{otherwise.} \end{cases}$$

With these f-image, the mapping $f : X \to Y$ is extended to the respective power sets by

$$\widehat{f} : \mathbb{P}(X) \to \mathbb{P}(Y), \quad A \mapsto f(A).$$

In the same vein, given a mapping $f : X \to Y$, it can be extended to the fuzzy power sets $F(X)$, $F(Y)$ by

$$\widehat{f} : F(X) \to F(Y)$$

$$\widehat{f}(\mu)(y) = \sup\{\mu(x); f(x) = y\}, \quad \text{for all } y \in Y,$$

and $\widehat{f}$ is known as the 'extension' of f to the fuzzy parts, and the definition as the Zadeh's *Extension Principle*.

For example, if $f : [0, 10] \to [0, 1]$, is given by $f(x) = 1 - \frac{x}{10}$, the fuzzy set $\mu(x) = \frac{x}{10}$ in $[0, 1]^{[0,10]}$ extends to the fuzzy set in $[0, 1]$, $\widehat{f}(\mu)(y) = \sup\{\mu(x); f(x) = y\} = \sup\{\frac{x}{10}; 1 - \frac{x}{10} = y\} = 1 - y$, for all $y \in [0, 1]$.

If $X = \{1, 2, 3, 4\}$, $Y = \{a, b, c\}$, the mapping $f : X \to Y$ such that

$$f(1) = f(2) = a, \ \ f(3) = f(4) = b,$$

extends the fuzzy set $\mu = 1/1 + 0.4/2 + 1/3 + 0.7/4$ in $F(X)$, to the fuzzy set $\widehat{f}(\mu)$ in $F(Y)$ with values,

$$\begin{aligned}
&\widehat{f}(\mu)(a) = \max\{\mu(x); x \in f^{-1}(a)\} = \max\{\mu(1), \mu(2)\} = \max(1, 0.4) = 1\\
&\widehat{f}(\mu)(b) = \max\{\mu(3), \mu(4)\} = 1\\
&\widehat{f}(\mu)(c) = 0, \text{ since } f^{-1}(c) = \varnothing.
\end{aligned}$$

Hence,

$$\widehat{f}(\mu) = 1/a + 1/b,$$

that corresponds to the crisp subset $\{a, b\}$ of Y.

Notice that if $\mu = \mu_A \in \{0, 1\}^X$, it is $\widehat{f}(\mu_A) = \mu_{f(A)}$, that is not only a crisp subset of Y, but coincides with the classical extension $f(A)$ of A. Nevertheless, as it is shown by the above example, it can happen that $\widehat{f}(\mu) \in \mathbb{P}(Y)$ with $\mu \in F(X) - \mathbb{P}(X)$.

2.1.3 Preservation of the Classical Case

Like with the cartesian product and with the extension principle, all operations with fuzzy sets must reproduce, when the data are crisp, the corresponding result obtained in the crisp theory. This is the *principle of preservation* of the classical case, that is forced by the will, and the necessity, of including all 'the classical' as a particular case of the algebras of fuzzy sets.

To illustrate this preservation's principle, let us show a negative example. With $X = [0, 1]$, and all $\mu \in [0, 1]^{[0,1]}$, the function

$$\mu^*(x) = 1 - \mu(1 - x),$$

verifies:

- $\mu_0^*(x) = 1 - \mu_0(1-x) = 1$: $\mu_0^* = \mu_1$
- $\mu_1^*(x) = 1 - \mu_1(1-x) = 1 - 1 = 0$: $\mu_1^* = \mu_0$
- $\mu \leqslant \sigma \Rightarrow 1 - \sigma(1-x) \leqslant 1 - \mu(1-x) \Rightarrow \sigma^* \leqslant \mu^*$
- $\mu^{**}(x) = 1 - \mu^*(1-x) = 1 - [1 - \mu(x)] = \mu(x)$: $\mu^{**} = \mu$.

Hence, *it could seem* that the function $\mu \mapsto \mu^*$ can be taken as a "strong negation" for the fuzzy sets in [0, 1], but it is not the case. Notice that if $\mu \in \{0, 1\}^{[0,1]}$, then it should be also $\mu^* \in \{0, 1\}^{[0,1]}$, that is, if μ represents a classical subset of [0, 1], also μ^* should represent not only a classical subset but precisely the complement of μ. But with $A = [0, \frac{1}{2}] \subset X$,

$$\mu_A(x) = \begin{cases} 1, & 0 \leqslant x \leqslant 0.5, \\ 0, & 0.5 < x \leqslant 1, \end{cases}$$

follows,

$$\mu_A^*(x) = 1 - \mu_A(1-x) = 1 - \begin{cases} 1, & 0.5 < x \leqslant 1, \\ 0, & 0 \leqslant x \leqslant 0.5, \end{cases} = \begin{cases} 0, & 0.5 < x \leqslant 1, \\ 1, & 0 \leqslant x \leqslant 0.5 \end{cases}$$

that represents the subset $[0, 0.5]$, but not $A^c = (0.5, 1]$. *The unary operation* $*$ *violates the preservation principle*, and hence it cannot be taken into account to negate fuzzy sets.

2.1.4 Resolution

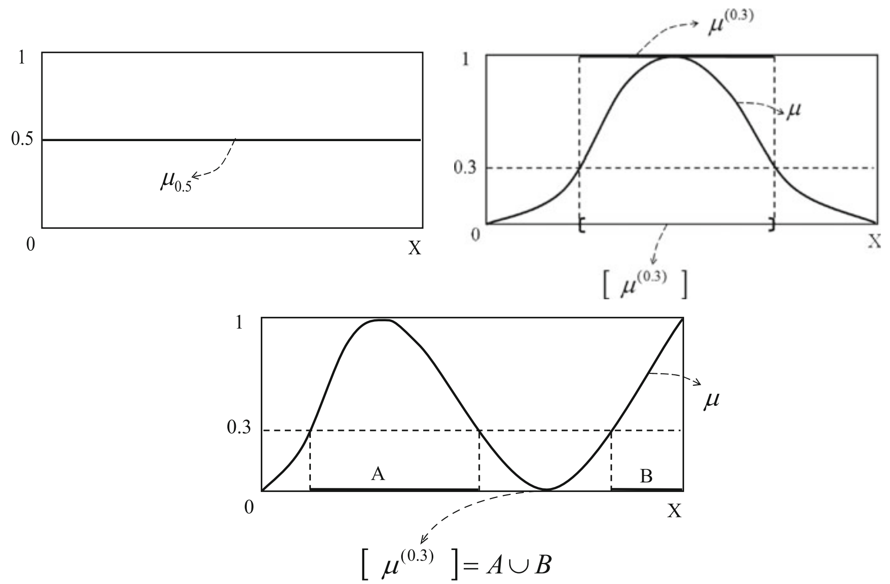
Let us denote by μ_r the constant fuzzy sets in $[0, 1]^X$, $\mu_r(x) = r$, for $r \in [0, 1]$ and all $x \in X$. Notice that in $\{0, 1\}^X$ there are only the "constants" μ_0 and μ_1, that correspond to the sets $\varnothing$ and X, respectively.

Given $\mu \in [0, 1]^X$, let us denote by $\mu^{(r)}$ the fuzzy (crisp) set

$$\mu^{(r)}(x) = \begin{cases} 1, & \text{if } r \leqslant \mu(x), \\ 0, & \text{otherwise}, \end{cases}$$

for all $r \in [0, 1]$, and by $[\mu^{(r)}]$ the corresponding classical subset $\{x \in X; r \leqslant \mu(x)\}$. These sets are called the *r-cuts* of μ and it is always $[\mu^{(0)}] = X$.

For example, in the following figures are shown, respectively, the constant fuzzy set $\mu_{0.5}$, and the 0.3-cut of two different fuzzy sets.



Notice that when $\mu_A \in \{0, 1\}^X$, it is

- $[\mu_A^{(0)}] = X$, since for all $x \in X$ it is $0 \leqslant \mu(x)$
- If $r > 0$, $[\mu_A^{(r)}] = A$, since for all $x \in A$ it is $0 < r \leqslant 1 = \mu_A(x)$,

then, *the only* r-cuts of a crisp subset A of X are X and A.

If $r \leqslant s$, since $s \leqslant \mu(x)$ implies $r \leqslant \mu(x)$, it results $[\mu^{(s)}] \subset [\mu^{(r)}]$: r-cuts are decreasing when their indices increase.

Let us show an example with $X = \{1, 2, 3, 4, 5\}$ and $\mu = 0.8/1 + 0.6/2 + 0.7/3 + 1/4 + 1/5$, where the only significative values for the r-cuts are 0.6, 0.7, 0.8, and 1:

- $[\mu^{(0.6)}] = \{1, 2, 3, 4, 5\} = X$
- $[\mu^{(0.7)}] = \{1, 3, 4, 5\}$
- $[\mu^{(0.8)}] = \{1, 4, 5\}$
- $[\mu^{(1)}] = \{4, 5\}$.

It is clear that $0.6 \leqslant 0.7 \leqslant 0.8 \leqslant 1$, and $[\mu^{(1)}] \subset [\mu^{(0.8)}] \subset [\mu^{(0.7)}] \subset [\mu^{(0.6)}]$.

Theorem 2.1.1 (Theorem of resolution) *For all* $\mu \in [0, 1]^X$, *is* $\mu(x) = \max\{r \in [0, 1]; \min(\mu_r(x), \mu^{(r)}(x)\}$, *for all* $x \in X$.

Proof

$$\max_{0 \leqslant r \leqslant 1} \min(\mu_r(x), \mu^{(r)}(x)) = \max_{0 \leqslant r \leqslant 1} \min(r, \mu^{(r)}(x))$$

$$= \max_{0 \leqslant r \leqslant 1} \min\left(r, \begin{cases} 1, & \text{if } r \leqslant \mu(x), \\ 0, & \text{otherwise,} \end{cases}\right) = \max_{0 \leqslant r \leqslant 1} \begin{cases} r, & \text{if } r \leqslant \mu(x), \\ 0, & \text{otherwise,} \end{cases} = \mu(x).$$

□

Example 2.1.2 With X and μ in the last example, it is

$$\mu(x) = \max(\min(0.6, \mu^{(0.6)}(x)), \min(0.7, \mu^{(0.7)}(x)), \min(0.8, \mu^{(0.8)}(x)), \min(1, \mu^{(1)}(x))),$$

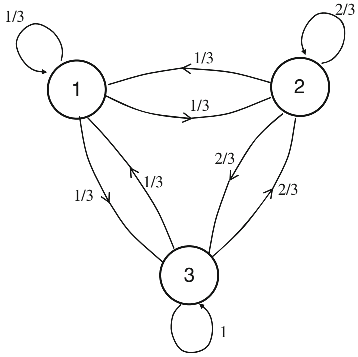
and, for instance,

$\mu(1) = \max(\min(0.6, 1), \min(0.7, 1), \min(0.8, 1), \min(1, 0)) = 0.8$
$\mu(4) = \max(\min(0.6, 1), \min(0.7, 1), \min(0.8, 1), \min(1, 1)) = 1$
etc.

Example 2.1.3 With $X = \{1, 2, 3\}$, take $\mu : X \times X \to [0, 1]$, given by $\mu(i, j) = \frac{\min(i,j)}{3}$. This fuzzy set in $X \times X$ can be represented either by the matrix

$$\begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} & 1 \end{pmatrix},$$

or by the graph



Since the matrices of $\mu^{(1)}$, $\mu^{(\frac{2}{3})}$, and $\mu^{(\frac{1}{3})}$ are respectively

$$\begin{pmatrix} 0&0&0 \\ 0&0&0 \\ 0&0&1 \end{pmatrix}, \begin{pmatrix} 0&0&0 \\ 0&1&1 \\ 0&1&1 \end{pmatrix}, \begin{pmatrix} 1&1&1 \\ 1&1&1 \\ 1&1&1 \end{pmatrix},$$

it results

$$\max\left(\min\left(1, \begin{pmatrix} 0&0&0 \\ 0&0&0 \\ 0&0&1 \end{pmatrix}\right), \min\left(\frac{2}{3}, \begin{pmatrix} 0&0&0 \\ 0&1&1 \\ 0&1&1 \end{pmatrix}\right), \min\left(\frac{1}{3}, \begin{pmatrix} 1&1&1 \\ 1&1&1 \\ 1&1&1 \end{pmatrix}\right)\right)$$

$$= \max\left(\begin{pmatrix} 0&0&0 \\ 0&0&0 \\ 0&0&1 \end{pmatrix}, \begin{pmatrix} 0&0&0 \\ 0&\frac{2}{3}&\frac{2}{3} \\ 0&\frac{2}{3}&\frac{2}{3} \end{pmatrix}, \begin{pmatrix} \frac{1}{3}&\frac{1}{3}&\frac{1}{3} \\ \frac{1}{3}&\frac{1}{3}&\frac{1}{3} \\ \frac{1}{3}&\frac{1}{3}&\frac{1}{3} \end{pmatrix}\right) = \begin{pmatrix} \frac{1}{3}&\frac{1}{3}&\frac{1}{3} \\ \frac{1}{3}&\frac{2}{3}&\frac{2}{3} \\ \frac{1}{3}&\frac{2}{3}&1 \end{pmatrix},$$

accordingly with the theorem of resolution.

2.2 The Concept of an 'Algebra of Fuzzy Sets'

2.2.1 Introduction

Functions

$$\mu \in F(X) = [0, 1]^X,$$

will be labeled only when it is some predicate P in X such that $\mu_P = \mu$, and it is obvious that it could be the fact of having $\mu_P = \mu_Q = \mu_R \cdots = \mu$, in which case the predicates $P, Q, R, \ldots$ are exact synonyms in X. Notwithstanding there are much more functions in $[0, 1]^X$ than predicates in X, and given a not previously labeled $\mu \in [0, 1]^X$, it can be 'artificially' introduced the predicate M $(= \mu)$ such that,

$$\text{Degree up to which } x \text{ is } M = \mu(x), \quad \text{for all } x \text{ in } X.$$

- Notice that $F(X)$ will be taken as 'ordered' (partially) by means of the binary pointwise relation

$$\mu \leqslant \sigma \Leftrightarrow \mu(x) \leqslant \sigma(x), \quad \text{for all } x \in X,$$

that induces the pointwise identity

$$\mu = \sigma \Leftrightarrow \mu \leqslant \sigma \text{ and } \sigma \leqslant \mu \Leftrightarrow \mu(x) = \sigma(x), \quad \text{for all } x \in X.$$

The pointwise relation $\leqslant$ is also called the 'inclusion', and $\mu \leqslant \sigma$ denoted by 'μ is included in σ'. It enjoys the laws reflexive, antisymmetric and transitive.

It will be always considered that $F(X)$ denotes, at least, the structure $([0, 1]^X; \leqslant; =)$. Observe that if $\mu_A, \mu_B \in \{0, 1\}^X$, that is, A and B are in $\mathbb{P}(X)$, then it follows

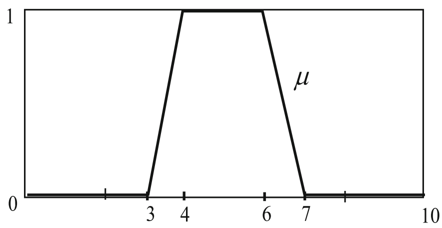
$$\mu_A \leqslant \mu_B \Leftrightarrow A \subset B; \quad \mu_A = \mu_B \Leftrightarrow A = B,$$

and

$$x \in A \Leftrightarrow \mu_A(x) = 1; \quad x \notin A \Leftrightarrow \mu_A(x) = 0.$$

The classical symbol $\in$ is the fuzzy symbol $\underset{1}{\in}$, and $\notin$ is $\underset{0}{\in}$.

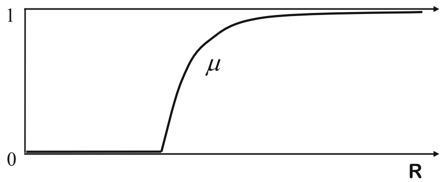
For example, with the fuzzy set $\underset{\sim}{P}$ given by function μ in the next figure it is $x \notin \underset{\sim}{P}$ if $0 \leqslant x \leqslant 3$, and $7 \leqslant x \leqslant 10$, but $x \in \underset{\sim}{P}$ if $4 \leqslant x \leqslant 6$, and $x \underset{\mu(x)}{\in} \underset{\sim}{P}$, if $x \in (3, 4) \cup (6, 7)$ with $0 < \mu(x) < 1$. If $x = 3.5$, since it is $\mu(x) = x - 3$, when $x \in (3, 4)$, it is $3.5 \underset{0.5}{\in} \underset{\sim}{P}$.



- The height of $\mu \in F(X)$ is $H(\mu) = \underset{x\in X}{Sup}\, \mu(x) = \sup \mu$. In the last example, it is $H(\mu) = 1$. In the finite example

$$\mu = 0.7/x_1 + 0.9/x_2 + 0.7/x_3,$$

in $X = \{x_1, x_2, x_3, x_4\}$, it is $H(\mu) = 0.9$. In a case like



it is $H(\mu) = 1$, although there is not any $x \in \mathbb{R}$ such that $\mu(x) = 1$. If there is some $x \in X$ such that $\mu(x) = 1$, it is said that μ is a *normalized* fuzzy set.
- In the case X is finite, $X = \{x_1, \ldots, x_n\}$, the number

$$|\mu| = \sum_{i=1}^{n} \mu(x_i)$$

is the *crisp-cardinality*, or sigma-count, of μ, a name coming from the fact that if $A \subseteq X$ has p elements it is $\sum_{i=1}^{n} \mu(x_i) = p$. Obviously, $\mu_\varnothing = \mu_0$, gives $|\mu_0| = 0$, $\mu_X = \mu_1$, gives $|\mu_1| = n$, and, $\mu \leqslant \sigma$ implies $|\mu| \leqslant |\sigma|$.

Remark 2.2.1 The pointwise definition of fuzzy sets inclusion implies that, for example, the fuzzy sets

$$\mu = 0.7/x_1 + 0.8/x_2 + 1/x_3 + 0.7/x_4$$
$$\sigma = 0.70001/x_1 + 0.7/x_2 + 1/x_3 + 0.6/x_4$$

in X, do not verify $\mu \leqslant \sigma$ although it is $\sigma(x_2) < \mu(x_2)$, $\sigma(x_3) = \mu(x_3)$, $\sigma(x_4) < \mu(x_4)$, but $\sigma(x_1) > \mu(x_1)$, with $\sigma(x_1) - \mu(x_1) = 0.00001$. Pointwise 'inclusion' is strongly affected by very small variations of the membership values. Actually, it is not a flexible, or fuzzy, concept, but a crisp one.

Because of this, it could be preferable to take the inclusion of fuzzy sets as an gradable concept $\leqslant_r$ $(r \in [0, 1])$, and a used definition of which is

$$\mu \ \leqslant_r \ \sigma \Leftrightarrow \frac{|\min(\mu, \sigma)|}{|\mu|} \leqslant r,$$

with $|\min(\mu, \sigma)| = \sum_{i=1}^{n} \min(\mu(x_i), \sigma(x_i))$.

In last example, it is $|\min(\mu, \sigma)| = 0.7 + 0.7 + 1 + 0.6 = 3$, $|\mu| = 0.7 + 0.8 + 1 + 0.7 = 3.2$, and $r = 3/3.2 = 0.9375 \approx 0.94$. That is, $\mu \underset{0.9375}{\leqslant} \sigma : \mu$ 'is almost included in' σ.

Since $|\sigma| = 0.70001 + 0.7 + 1 + 0.6 = 3.00001$, it is $\frac{|\min(\mu,\sigma)|}{|\sigma|} = 0.9999$, or $\sigma \leqslant_{0.9999} \mu$. That is, σ is more included in μ, than μ is included in σ!

Remark 2.2.2 Of course, if $\mu \leqslant \sigma$, it is $\min(\mu, \sigma) = \mu$, and $r = 1$, that is,

$$\mu \leqslant \sigma \Rightarrow \mu \underset{1}{\leqslant} \sigma.$$

Nevertheless, since it is only

$$\sum \min(\mu(x_i), \sigma(x_i)) \leqslant \min(\sum \mu(x_i), \sum \sigma(x_i)),$$

from $\mu \underset{1}{\leqslant} \sigma$ (or $|\min(\mu, \sigma)| \leqslant |\mu|$) it does not necessarily follow $\mu \leqslant \sigma$.

Let us show an example with crisp subsets. If $X = \{1, 2, 3, 4, 5, 6, 7\}$, and $A = \{1, 3, 5, 7\}$, $B = \{1, 3, 5, 6\}$, it is

$$\sum_{i=1}^{7} \min(\mu_A(i), \mu_B(i)) = 3, \sum_{i=1}^{7} \mu_B(i) = 4, \sum_{i=1}^{7} \mu_A(i) = 4.$$

hence

$$\mu_A \leqslant_{\frac{3}{4}} \mu_B, \text{ or } A \subset_{\frac{3}{4}} B$$

$$\mu_B \leqslant_{\frac{3}{4}} \mu_A, \text{ or } B \subset_{\frac{3}{4}} A$$

2.2.2 *Algebras of Fuzzy Sets*

Once $F(X) = ([0, 1]^X; \leqslant; =)$ is taken, a *general algebra of fuzzy sets* comes from endowing $F(X)$ with three operations:

1. $\prime : [0, 1]^X \to [0, 1]^X$,
2. $\cdot : [0, 1]^X \times [0, 1]^X \to [0, 1]^X$,
3. $+ : [0, 1]^X \times [0, 1]^X \to [0, 1]^X$,

respectively called the *complement* μ' of μ, the *intersection* $\mu \cdot \sigma$ of 'μ and σ', and the *union* $\mu + \sigma$ of 'μ or σ'. Then $([0, 1]^X; \leqslant; =; \ \cdot; +;')$, is called an *algebra of fuzzy sets*, provided the following laws do hold:

(a) If $\mu \leqslant \sigma$, then $\gamma \cdot \mu \leqslant \gamma \cdot \sigma$, and $\mu \cdot \gamma \leqslant \sigma \cdot \gamma$, for all $\gamma \in [0, 1]^X$
(b) If $\mu \leqslant \sigma$, then $\mu + \gamma \leqslant \sigma + \gamma$, and $\gamma + \mu \leqslant \gamma + \sigma$ for all $\gamma \in [0, 1]^X$
(c) If $\mu \leqslant \sigma$, then $\sigma' \leqslant \mu'$
(d) For any $\mu \in [0, 1]^X$, $\mu \cdot \mu_1 = \mu_1 \cdot \mu = \mu$, $\mu + \mu_0 = \mu_0 + \mu = \mu$
(e) For all $\mu_A, \mu_B \in \{0, 1\}^X$, $\mu'_A = \mu_{A^c}$, $\mu_A \cdot \mu_B = \mu_{A\cap B}$, $\mu_A + \mu_B = \mu_{A\cup B}$ (preservation of the classical case).

Remark 2.2.3 It is not difficult to prove that no general algebra of fuzzy sets is a Boolean algebra. The proof comes from the fact that to be a Boolean algebra would imply $\mu \cdot \mu' = \mu_0$ and $\mu + \mu' = \mu_1$ for all $\mu \in [0, 1]^X$, and consists in finding some μ for which these equalities are not satisfied.

Remark 2.2.4 It is immediate to prove that $\mu \cdot \mu_0 = \mu_0 \cdot \mu = \mu_0$, $\mu + \mu_1 = \mu_1 + \mu = \mu_1$ for all $\mu \in [0, 1]^X$

Remark 2.2.5 Notice that the laws $\mu \cdot \sigma = \sigma \cdot \mu$ (commutativity of the intersection), $\mu + \sigma = \sigma + \mu$ (commutativity of the union), and $\mu'' = \mu$ (involution of the complement) are not supposed to be always verified. Nor it is supposed that the algebras ($[0, 1]^X \cdot, +, \prime$) *are dual ones*, that is, the so-called De Morgan laws,

$$(\mu + \sigma)' = \mu' \cdot \sigma', (\mu \cdot \sigma)' = \mu' + \sigma',$$

are not supposed to hold in general. It is neither supported $\mu \cdot \mu = \mu$, and $\sigma + \sigma = \sigma$, nor $\mu \cdot (\sigma \cdot \lambda) = (\mu \cdot \sigma) \cdot \lambda$ and $\mu + (\sigma + \lambda) = (\mu + \sigma) + \lambda$, nor that $\cdot$ and $+$ are associative, $\mu \cdot (\sigma \cdot \lambda) = (\mu \cdot \sigma) \cdot \lambda$, and $\mu + (\sigma + \lambda) = (\mu + \sigma) + \lambda$.

Theorem 2.2.6 *For any* $\mu, \sigma \in [0, 1]^X : \mu \cdot \sigma \leqslant \min(\mu, \sigma) \leqslant \max(\mu, \sigma) \leqslant \mu + \sigma$.

Proof From $\mu \leqslant \mu_1$ ($\mu(x) \leqslant 1$, for all $x \in X$), follows $\mu \cdot \sigma \leqslant \mu_1 \cdot \sigma = \sigma$. From $\sigma \leqslant \mu_1$, follows $\mu \cdot \sigma \leqslant \mu_1 \cdot \mu = \mu$. Thus,

$$\begin{aligned}&(\mu \cdot \sigma)(x) \leqslant \sigma(x), (\mu \cdot \sigma)(x) \leqslant \mu(x)\\ &\quad \Rightarrow (\mu \cdot \sigma)(x) \leqslant \min(\mu(x), \sigma(x)) = \min(\mu, \sigma)(x),\end{aligned}$$

or $\mu \cdot \sigma \leqslant \min(\mu, \sigma)$. Hence, the operation min is the greatest possible intersection of fuzzy sets. Analogously, $\mu_0 \leqslant \mu, \mu_0 \leqslant \sigma \Rightarrow \mu_0 + \sigma = \sigma \leqslant \mu + \sigma, \mu + \mu_0 = \mu \leqslant \mu + \sigma$, and $\max(\mu, \sigma) \leqslant \mu + \sigma$: The operation *max* is the smallest possible union of fuzzy sets. □

Obviously, for all $\mu, \sigma \in [0, 1]^X$: $\mu \cdot \sigma \leqslant \mu \leqslant \mu + \sigma$, $\mu \cdot \sigma \leqslant \sigma \leqslant \mu + \sigma$.

Theorem 2.2.7 *An operation* $* : [0, 1]^X \times [0, 1]^X \to [0, 1]^X$ *is called idempotent if and only if* $\mu * \mu = \mu$, *for all* $\mu \in [0, 1]^X$

- *The intersection* $\cdot$ *is idempotent if and only if* $\cdot = \min$
- *The union* $+$ *is idempotent if and only if* $+ = \max$

Proof The operations min, and max are obviously idempotent. Let us show that if $\cdot$ is idempotent, it must be $\cdot = \min$. By Theorem 2.2.6, it is always $\mu \cdot \sigma \leqslant \min(\mu, \sigma)$, and the idempotency of $\cdot$ implies

$$\min(\mu, \sigma) \cdot \min(\mu, \sigma) = \min(\mu, \sigma).$$

But $\min(\mu, \sigma) \leqslant \mu$, $\min(\mu, \sigma) \leqslant \sigma$, imply $\min(\mu, \sigma) \cdot \min(\mu, \sigma) \leqslant \mu \cdot \sigma$, that is

$$\min(\mu, \sigma) \leqslant \mu \cdot \sigma,$$

and, by Theorem 2.2.6, $\min(\mu, \sigma) = \mu \cdot \sigma$. A similar proof applies to $+$ and *max*. □

Theorem 2.2.8 (Absortion laws)

- $\mu \cdot (\mu + \sigma) = \mu$ *holds for all* $\mu, \sigma \in [0, 1]^X \Leftrightarrow \cdot = \min$
- $\mu + (\mu \cdot \sigma) = \mu$ *holds for all* $\mu, \sigma \in [0, 1]^X \Leftrightarrow + = \max$

Proof If $\cdot = \min$, the formula $\min(\mu, \mu + \sigma) = \mu$ does hold, since $\mu \leqslant \mu + \sigma$. Provided it is always $\mu \cdot (\mu + \sigma) = \mu$, taking $\sigma = \mu_0$ follows $\mu \cdot (\mu + \mu_0) = \mu \cdot \mu = \mu$, that holds if and only if $\cdot = \min$. If $+ = \max$, the formula $\max(\mu, \mu \cdot \sigma) = \mu$ does hold since $\mu \cdot \sigma \leqslant \mu$. Provided it is always $\mu + (\mu \cdot \sigma) = \mu$, taking $\sigma = \mu_1$ follows $\mu + (\mu \cdot \mu_1) = \mu + \mu = \mu$, that holds if and only if $+ = \max$. □

Theorem 2.2.9 (Duality, or De Morgan laws) *Provided the* complement $'$ is involutive ($(\mu')' = \mu'' = \mu$, *for all* $\mu \in [0, 1]^X$), *the algebra of fuzzy sets* ($[0, 1]^X$, min, max, $'$) *is a dual algebra.*

Proof If $'$ is involutive, from $\mu' + \sigma' = (\mu \cdot \sigma)'$ it follows $\mu' \cdot \sigma' = (\mu + \sigma)'$ since $\mu + \sigma = \mu'' + \sigma'' = (\mu' \cdot \sigma')'$, hence $(\mu + \sigma)' = \mu' \cdot \sigma'$. The converse is proven in the same way. And the two De Morgan laws

$$(\mu \cdot \sigma)' = \mu' + \sigma', \quad (\mu + \sigma)' = \mu' \cdot \sigma'$$

result equivalent. Then, it is enough to prove $\max(\mu, \sigma) = (\min(\mu', \sigma'))'$, for all μ, σ in $[0, 1]^X$. Since

$$\min(\mu', \sigma') \leqslant \mu', \quad \min(\mu', \sigma') \leqslant \sigma',$$

it follows

$$\mu = \mu'' \leqslant (\min(\mu', \sigma'))', \quad \sigma = \sigma'' \leqslant (\min(\mu', \sigma'))',$$

and

$$\max(\mu, \sigma) \leqslant (\min(\mu', \sigma'))' \tag{2.1}$$

On the other hand,

$$\mu \leqslant \max(\mu, \sigma) \Rightarrow (\max(\mu, \sigma))' \leqslant \mu'$$

$$\sigma \leqslant \max(\mu, \sigma) \Rightarrow (\max(\mu, \sigma))' \leqslant \sigma'$$

imply $(\max(\mu, \sigma))' \leqslant \min(\mu', \sigma')$, or

$$(\min(\mu', \sigma'))' \leqslant \max(\mu, \sigma). \tag{2.2}$$

Now, from (2.2) and (2.1), follows the result. □

Theorem 2.2.10 (Kleene's Law) *In all general algebra of fuzzy sets it holds the law*

$$\mu \cdot \mu' \leqslant \sigma + \sigma',$$

for all μ, σ *in* $[0, 1]^X$.

Proof We have to prove that, for any $x \in X$, it is $(\mu \cdot \mu')(x) \leqslant (\sigma + \sigma')(x)$. But it only can be either $\mu(x) \leqslant \sigma(x)$, or $\sigma(x) \leqslant \mu(x)$ for each $x \in X$. In the firs case, it is $(\mu \cdot \mu')(x) \leqslant \min(\mu(x), \mu'(x)) \leqslant \mu(x) \leqslant \sigma(x) \leqslant \max(\sigma(x), \sigma'(x)) = (\max(\sigma, \sigma')(x) \leqslant (\sigma + \sigma')(x)$. In the second case, it is $\mu'(x) \leqslant \sigma'(x)$, and $(\mu \cdot \mu')(x) \leqslant \min(\mu(x), \mu'(x)) \leqslant \mu'(x) \leqslant \sigma'(x) \leqslant \max(\sigma(x), \sigma'(x)) = (\max(\sigma, \sigma')(x) \leqslant (\sigma + \sigma')(x)$. Notice that provided μ and σ were crisp sets, the Kleene's law is reduced to $\mu_0 \leqslant \mu_1$. □

Remark 2.2.11 Concerning duality, Theorem 2.2.9 only states that the algebra given by the triplets (min, max, $'$), with $'$ involutive, are dual algebras. But they are not the only dual algebras. For example, with $\cdot$ = product,

$$(\mu \cdot \sigma)(x) = \mu(x) \cdot \sigma(x), \quad \forall x \in X,$$

it is easy to proof that taking $\mu'(x) = 1 - \mu(x)$, and

$$(\mu + \sigma)(x) = 1 - (1 - \mu(x))(1 - \sigma(x)) = \mu(x) + \sigma(x) - \mu(x) \cdot \sigma(x),$$

it is $([0, 1]^X, \cdot, +, ')$ an algebra of fuzzy sets that since it is

$$\mu + \sigma = (\mu' \cdot \sigma')',$$

is a dual algebra. Nevertheless, with $\mu'(x) = 1 - \mu(x)$, $(\mu \cdot \sigma)(x) = \mu(x) \cdot \sigma(x)$, and $(\mu + \sigma)(x) = \max(\mu(x), \sigma(x))$, we get an algebra that is not dual since

$$(\mu' \cdot \sigma')'(x) = \mu(x) + \sigma(x) - \mu(x) \cdot \sigma(x)$$

does not coincides with $\max(\mu(x), \sigma(x))$, as it is easy to see.

Remark 2.2.12 It is easy to prove that, for each algebra of fuzzy sets $([0, 1]^X, \cdot, +, ')$, the operation

$$\mu +' \sigma = (\mu' \cdot \sigma')',$$

gives the new algebra $([0, 1]^X, \cdot, +', ')$. If the complement $'$ is involutive ($\mu'' = \mu$), then $(\mu +' \sigma)' = \mu' \cdot \sigma'$.

Analogously, with the operation $\mu \cdot' \sigma = (\mu +' \sigma')'$, one has the new algebra $([0, 1]^X, \cdot', +', ')$ and, if $'$ is involutive, $(\mu \cdot' \sigma)' = \mu +' \sigma'$.

2.2.3 Non-contradiction and Excluded-Middle

An statement is self-contradictory whenever entails its negation. For example, the only classical set that is self-contradictory is the empty one:

$$A \subseteq A^c \Rightarrow A \cap A \subseteq A \cap A^c \Rightarrow A \subseteq \varnothing \Rightarrow A = \varnothing.$$

Perhaps, this is the reason of the difficulty children do have on accepting that $\varnothing$ is a set!

Within an algebra of fuzzy sets there are many self-contradictory fuzzy sets. For example, with $N = 1 - \mathrm{id}$ it is

$$\mu \leqslant \mu' \Leftrightarrow \mu(x) \leqslant 1 - \mu(x) \Leftrightarrow \mu(x) \leqslant \frac{1}{2}, \quad \forall x \in X,$$

hence: μ is self-contradictory if and only if $\mu \leqslant \mu_{\frac{1}{2}}$. Analogously, with the strong negation $N(x) = \frac{1-x}{1+x}$, it is

$$\mu \leqslant \mu' \Leftrightarrow \mu(x) \leqslant \frac{1-\mu(x)}{1+\mu(x)} \Leftrightarrow \mu(x)^2 + 2\mu(x) - 1 \leqslant 0$$
$$\Leftrightarrow \mu(x) \leqslant \sqrt{2} - 1, \quad \forall x \in X,$$

that is, μ is self-contradictory if and only id $\mu \leqslant \mu_{\sqrt{2}-1}$.

Notice that 1/2 is the fixed-point of the strong negation $N = 1 - \mathrm{id}$ ($1 - n = n \Leftrightarrow n = 1/2$), and that $\sqrt{2} - 1$ is the fixed-point of the strong negation $N = \frac{1-id}{1+id}$ ($\frac{1-n}{1+n} = n \Leftrightarrow n = \sqrt{2} - 1$).

Notice that if $\mu_P \leqslant \mu_{aP}$, since it is always supposed that $\mu_{aP} \leqslant \mu_{not\,P}$, it follows $\mu_P \leqslant \mu_{not\,P}$, and μ_P is self-contradictory.

The classical principle of non-contradiction, "it is impossible to have both an statement and its negation", could be interpreted as "P and not P is impossible", or "P and not P is self-contradictory". All general algebras of fuzzy sets do verify the principle of non-contradiction once stated in this form.

Theorem 2.2.13 *If $([0, 1]^X, \cdot, +, ')$ is an algebra of fuzzy sets, it holds the principle of non-contradiction stated by: $\mu \cdot \mu' \leqslant (\mu \cdot \mu')'$ for all $\mu \in [0, 1]^X$. That is, for all $\mu \in [0, 1]^X$, $\mu \cdot \mu'$ is self-contradictory.*

Proof It is $\mu \cdot \mu' \leqslant \min(\mu, \mu') \leqslant \mu$, and $\mu \cdot \mu' \leqslant \min(\mu, \mu') \leqslant \mu'$; from the first inequality it follows $\mu' \leqslant (\mu \cdot \mu')'$, and then with the second follows $\mu \cdot \mu' \leqslant (\mu \cdot \mu')'$. □

Notice that no additional hypotheses on the connective $\cdot$, and the complement $'$, are needed for the proof of this theorem. In the algebras of fuzzy sets the non-contradiction principle is a theorem: the algebra's axioms imply the principle. It is not true, as it is sometimes stated, that fuzzy sets do not verify the principle of non-contradiction in which science is based.

The classical principle of Excluded-Middle, "It is always P or not P", can be interpreted as "Not (P or Not P) is a self-contradiction"and it is verified by all algebra of fuzzy sets.

Theorem 2.2.14 *If* $([0, 1]^X, \cdot, +, ')$ *is an algebra of fuzzy sets, it holds the principle of Excluded-Middle stated by*: $(\mu + \mu')' \leqslant ((\mu + \mu')')'$ *for all* $\mu \in [0, 1]^X$. *That is, for all* $\mu \in [0, 1]^X$, $(\mu + \mu')'$ *is self-contradictory.*

Proof It is,

- $\mu \leqslant \max(\mu, \mu') \leqslant \mu + \mu' \Rightarrow (\mu + \mu')' \leqslant \mu' \Rightarrow (\mu')' \leqslant ((\mu + \mu')')'$
- $\mu' \leqslant \max(\mu, \mu') \leqslant \mu + \mu' \Rightarrow (\mu + \mu')' \leqslant (\mu')'$

then $(\mu + \mu')' \leqslant ((\mu + \mu')')'$. □

Notice that no additional hypotheses on the connective $+$, and the complement $'$, are needed for the proof of this theorem: In the algebra of fuzzy sets the excluded-middle principle is a theorem. In conclusion,

In all algebra of fuzzy sets $([0, 1]^X, \cdot, +, ')$,

- *The logic principles of non-contradiction and excluded-middle are theorems, once stated through the concept of self-contradiction.*

A very different situation appears if these two principles are stated as it is currently done within logic and classical set theory, that is, by stating

- "P and not P" is false
- "P or not P" is true,

or,

- There is no x in X such that "x is P and x is not P"
- For all x in X it is "x is P or x is not P"

translated into

- $(\mu_P \cdot \mu_{not\ P})(x) = 0$, for all x in X
- $(\mu_P + \mu_{not\ P})(x) = 1$, for all x in X

that corresponds to "solve" the equations with fuzzy sets,

$$\mu \cdot \mu' = \mu_0, \ \mu + \mu' = \mu_1,$$

that is, to find for which intersections $\cdot$ and which unions $+$, these equations do hold.

Of course, they do not hold in all cases, for example, with $N = 1 - \text{id}$,

- If $\cdot = \min$, it is not always $\min(\mu(x), 1 - \mu(x)) = 0$,
- If $+ = \max$, it is not always $\max(\mu(x), 1 - \mu(x)) = 1$,
- If $\cdot = W$, $(W(a, b) = \max(0, a + b - 1))$ it is $W(a, 1 - a) = \max(0, a + 1 - a - 1)) = 0$, and $W(\mu(x), 1 - \mu(x))) = 0$ for all x in X
- If $+ = W^*$, $(W^*(a, b) = \min(1, a + b))$, it is $W^*(\mu(x), 1 - \mu(x))) = 1$ for all x in X.

That is, *there are algebras* of fuzzy sets *where this forms* of non-contradiction or excluded-middle *hold*, and *algebras where* this principles *do not jointly hold*. In the algebras with the triplet (min, max, $1 - id$) do not hold both principles, in the algebras with $(W, \max, 1 - id)$ it holds the principle of non-contradiction but not that of excluded-middle, in the algebras with $(\min, W^*, 1 - id)$ it holds the excluded-middle but not the principle of non-contradiction, and in the algebras with $(W, W^*, 1 - id)$ both principles hold.

Remark 2.2.15 Let us show that with $\mu \in \{0, 1\}^X$ it is always $\mu \cdot \mu' = \mu_0$ and $\mu + \mu' = \mu_1$. If $\mu \in \{0, 1\}^X$, denote $A = \{x \in X; \mu(x) = 1\}$. Obviously, $\mu = \mu_A$ and $\mu' = \mu_{A^c}$, hence

- $(\mu \cdot \mu') = \mu_A \cdot \mu_{A^c} = \mu_{A \cap A^c} = \mu_\varnothing = \mu_0$
- $(\mu + \mu') = \mu_A + \mu_{A^c} = \mu_{A \cup A^c} = \mu_X = \mu_1$.

Remark 2.2.16 Results in Theorems 2.2.13 and 2.2.14 challenge the usual statement that in fuzzy sets the basic principles of Non-contradiction and Excluded-middle fail. A statement that could conduct to believe that fuzzy set algebras are not properly grounded in a solid ground.

The fact is, notwithstanding, that these two principles were established before the current ways of considering the problems of logic and, of course, before the *nomenclature* of set theory. In set theory (or Boolean algebras), $A \cap A^c = \varnothing$ and $A \cap A^c \subset (A \cap A^c)^c$ are equivalent formulas since, as it was said, it is

$$B = \varnothing \Leftrightarrow B \subset B^c,$$

an equivalence only verified in the setting of ortholattices (of which Boolean algebras are a particular case), but that does not hold on weaker algebraic structures like it is, for example, the case of the above defined algebras of fuzzy sets. Let us call 'restricted' the principles stated by $\mu \cdot \mu' = \mu_0$ and $\mu + \mu' = \mu_1$.

2.2.4 Decomposable Algebras

Definition 2.2.17 An operation with fuzzy sets $* : [0, 1]^X \times [0, 1]^X \to [0, 1]^X$ is *decomposable, or functionally expressible*, if it exists a numerical operation $\widehat{*}$: $[0, 1] \times [0, 1] \to [0, 1]$, such that

$$(\mu * \sigma)(x) = \widehat{*}(\mu(x), \sigma(x)),$$

for all μ, σ in $[0, 1]^X$ and all x in X. Of course, by this formula, a numerical operation $\widehat{*}$ allows to define an operation $*$ for fuzzy sets.

For example, the operation min in $[0, 1]^X$ is decomposable since, by definition, it is

$$(\min(\mu, \sigma))(x) = \min(\mu(x), \sigma(x))$$

for all μ, σ in $[0, 1]^X$ and all x in X.

Definition 2.2.18 A function $f : [0, 1]^X \to [0, 1]^X$ is *decomposable, or functionally expressible*, if it exists a numerical function $\widehat{f} : [0, 1] \to [0, 1]$, such that

$$(f(\mu))(x) = \widehat{f}(\mu(x)),$$

for all μ in $[0, 1]^X$ and all x in X.

For example, the function $'$ defined by

$$\mu'(x) = 1 - \mu(x),$$

is decomposable because of $N_0 = 1 - \text{id}$ gives $\mu'(x) = N_0(\mu(x))$. With $X = [0, 1]$, the function defined by $\mu^*(x) = 1 - \mu(1 - x)$ is not decomposable, since if it were such, that is, if there is $N : [0, 1] \to [0, 1]$ such that $1 - \mu(1 - x) = N(\mu(x))$, it suffices to take $\mu, \sigma \in [0, 1]^X$ such that

- $\mu(0) = 0, \mu(1) = 1 \Rightarrow N(0) = 1 - \mu(1 - 0) = 1 - \mu(1) = 0, N(1) = 1 - \mu(1 - 1) = 1 - \mu(0) = 1$
- $\sigma(0) = 1, \sigma(1) = 0 \Rightarrow N(0) = 1 - \sigma(1 - 0) = 1 - \sigma(1) = 1, N(1) = 1 - \sigma(1 - 1) = 1 - \sigma(0) = 0$

that is absurd.

The algebras of fuzzy sets $([0, 1]^X, \cdot, +, ')$, can be

Decomposable	if the three operation $\cdot, +, '$ are decomposable
Partially decomposable	if at least one of the three operations $\cdot, +, '$ is decomposable
Non decomposable	if no one of the three operations is decomposable

In what follows we will only deal with decomposable algebras, that is, such that:

There are three functions $F : [0, 1]\times[0, 1] \to [0, 1]$, $G : [0, 1]\times[0, 1] \to [0, 1]$, and $N : [0, 1] \to [0, 1]$ with which

$$(\mu \cdot \sigma)(x) = F(\mu(x), \sigma(x)), (\mu + \sigma)(x) = G(\mu(x), \sigma(x)), \mu'(x) = N(\mu(x)),$$

for all $\mu, \sigma \in [0, 1]^X$, and all $x \in X$. For short, $\mu \cdot \sigma = F \circ (\mu \times \sigma)$, $\mu + \sigma = G \circ (\mu \times \sigma)$, $\mu' = N \circ \mu$.

In these cases, instead of $([0, 1]^X, \cdot, +, ')$ it is written $([0, 1]^X, F, G, N)$.

The laws verified by the triplet $(\cdot, +, ')$ force analogous laws for the triplet (F, G, N). For example, linked to the axioms, it is

(a) If $a \leqslant b$, then $F(a, c) \leqslant F(b, c)$, $F(c, a) \leqslant F(c, b)$, for all $c \in [0, 1]$
(b) If $a \leqslant b$, then $G(a, c) \leqslant G(b, c)$, $G(c, a) \leqslant G(c, b)$, for all $c \in [0, 1]$
(c) If $a \leqslant b$, then $N(b) \leqslant N(a)$
(d) $F(1, a) = F(a, 1) = a$, $G(0, a) = G(a, 0) = a$
(e) $N(0) = 1$, $N(1) = 0$, $F(0, a) = F(a, 0) = 0$, $G(1, a) = G(a, 1) = 1$

and linked to the Theorems 2.2.6–2.2.13 it is

- $F(a, b) \leqslant \min(a, b) \leqslant \max(a, b) \leqslant G(a, b)$ for all a, b in $[0, 1]$. In particular, it is $F \leqslant G$.
- F is idempotent ($F(a, a) = a$, for all a in $[0, 1]$), if and only if $F = \min$.
- G is idempotent ($G(a, a) = a$, for all a in $[0, 1]$), if and only if $G = \max$.
- It is $F(a, G(a, b)) = a$, for all a, b in $[0, 1]$, if and only if $F = \min$.
- It is $G(a, F(a, b)) = a$, for all a, b in $[0, 1]$, if and only if $G = \max$.
- A triplet (F, G, N) is called dual, or De Morgan triplet, if $F = N \circ G \circ (N \times N)$, or, $G = N \circ F \circ (N \times N)$, that is, $F(a, b) = N(G(N(a), N(b))$, or $G(a, b) = N(F(N(a), N(b))$, for all a, b in $[0, 1]$. Notice that, in this case, it is enough to know N and F to have G, or N and G to have F.
- If N is involutive ($N(N(a)) = a$, for all $a \in [0, 1]$ or $N^2 = N$), the triplet $(\min, \max, N)$ is a dual one.
- All triplet (F, G, N) verifies $F(a, N(a)) \leqslant G(b, N(b))$, for all a, b in $[0, 1]$,
- It is $F(a, N(a)) \leqslant N(F(a, N(a)))$, for all a in $[0, 1]$
- It is $N(F(a, N(a))) \leqslant N(N(F(a, N(a))))$, for all a in $[0, 1]$
- Given N involutive and F, and denoting by G_N the dual of F respect to N, $G_N(a, b) = N(F(N(a), N(b)))$, it follows $G_N(N(a), a) \leqslant N(G_N(N(a), a))$, that with
$$\mu + \sigma = G_N \circ (\mu \times \sigma),$$
gives $\mu' + \mu \leqslant (\mu' + \mu)'$.
- It is always $F(a, N(a)) \leqslant G(b, N(b))$, for all a, b in $[0, 1]$.

Remark 2.2.19 With classical sets, from $A \cap B \subseteq A \cup B$ it results $(A\cup B)\cup(A\cap B) = A\cup B$. In a decomposable theory, to have the analogous law $(\mu+\sigma)+(\mu \cdot \sigma) = \mu+\sigma$ it should be $G(G(a, b), F(a, b)) = G(a, b)$, for all a, b in $[0, 1]$, that is verified if, for example, $G = \max$, $F = \min$, or $G = W^*$, $F = W$

Remark 2.2.20 Let us show an example of a 'union' that is not-decomposable. Define

$$(\mu + \sigma)(x) = \begin{cases} \max(\mu(x), \sigma(x)), & \text{if } \mu \text{ or } \sigma \text{ are in} \{0, 1\}^X \text{ (crisp)} \\ \max(H(\mu), H(\sigma)), & \text{otherwise.} \end{cases}$$

It is easy to show that this operation verifies the laws b, d and e in Sect. 2.1.1. Hence, it is a union for fuzzy sets that, in addition, is commutative. It is not idempotent, since if $\mu \in [0, 1]^X - \{0, 1\}^X$, it is $(\mu + \mu)(x) = H(\mu) \neq \mu(x)$. It does not exist a function $G : [0, 1] \times [0, 1] \to [0, 1]$ such that

$$(\mu + \sigma)(x) = G(\mu(x), \sigma(x))$$

for all $x \in X$ and all $\mu, \sigma[0, 1]^X$. Indeed, let us suppose that such a G does exist, and take $\mu = \mu_{0.5}$. Then

- With $\sigma = \mu_0$, it is $(\mu + \sigma)(x) = \max(\frac{1}{2}, 0) = \frac{1}{2}$. Hence $G(\frac{1}{2}, 0) = \frac{1}{2}$.
- With $\sigma(x) = x$, is $(\mu + \sigma)(x) = \max(H(\mu), H(\sigma)) = \max(\frac{1}{2}, 1) = 1$, and $(\mu + \sigma)(0) = 1 = G(\frac{1}{2}, 0)$, that is absurd.
  To have a not-decomposable 'intersection', it is enough to define, with $\mu' = 1 - \mu$, the dual operation,

$$\begin{aligned} (\mu \cdot \sigma)(x) &= [(\mu' + \sigma')]'(x) = 1 - (\mu' + \sigma')(x) \\ &= \begin{cases} \min(\mu(x), \sigma(x)), & \text{if } \mu \text{ or } \sigma \text{ are in} \{0, 1\}^X \\ \max(H(\mu'), H(\sigma')), & \text{otherwise.} \end{cases} \end{aligned}$$

2.2.5 Standard Algebras of Fuzzy Sets

An standard algebra of fuzzy sets is a decomposable algebra of fuzzy sets such that:

1. $\mu \cdot \sigma = \sigma \cdot \mu$, for all μ, σ in $[0, 1]^X$ ($\cdot$ is commutative)
2. $\mu + \sigma = \sigma + \mu$, for all μ, σ in $[0, 1]^X$ ($+$ is commutative)
3. $\mu \cdot (\sigma \cdot \lambda) = (\mu \cdot \sigma) \cdot \lambda$, for all μ, σ, λ in $[0, 1]^X$ ($\cdot$ is associative)
4. $\mu + (\sigma + \lambda) = (\mu + \sigma) + \lambda$, for all μ, σ, λ in $[0, 1]^X$ ($+$ is associative)
5. $\mu'' = \mu$, for all μ in $[0, 1]^X$ ($'$ is involutive).

Hence, writing

$$\mu \cdot \sigma = T \circ (\mu \times \sigma), \;\; \mu + \sigma = S \circ (\mu \times \sigma), \;\; \mu' = N \circ \mu,$$

functions $T, S : [0, 1] \times [0, 1] \to [0, 1]$ and $N : [0, 1] \to [0, 1]$, in addition to the corresponding general properties stated before, must verify

- T is commutative, S is commutative
- T is associative, S is associative
- N is involutive,

that is:

- $T(a, b) = T(b, a)$, $S(a, b) = S(b, a)$, for all a, b in $[0, 1]$
- $T(a, T(b, c)) = T(T(a, b), c)$, $S(a, S(b, c)) = S(S(a, b), c)$, for all a, b, c in $[0, 1]$
- $N(N(a)) = a$, for all a in $[0, 1]$, or $N \circ N = \mathrm{id}$, or $N = N^{-1}$.

Functions T and S are called t-norms and t-conorms, respectively. Functions N are strong negations. Hence, $([0, 1], T, \leqslant)$ is an ordered semigroup with neutral 1, and absorbent 0, and $([0, 1], S, \leqslant)$ is also an ordered semigroup but with neutral 0 and absorbent 1. Since $N(1) = 0$, it seems that this two kind of ordered semigroups should show some character of duality. This duality goes in the way:

- If T is a t-norm, $T_N = N \circ S \circ (N \times N)$ is a t-conorm
- If S is a t-conorm, $S_N = N \circ S \circ (N \times N)$ is a t-norm

that are easy to prove. Of course, from Sect. 2.1.4,

- If T is a t-norm, $T \leqslant \min$, and min is a t-norm
- If S is a t-conorm, $\max \leqslant S$, and *max* is a t-conorm

Hence, for all t-norm T and all t-conorm S:

$$T \leqslant \min \leqslant \max \leqslant S,$$

in particular, $T \leqslant S$.[1] Even more, the function

$$Z(a, b) = \begin{cases} b, & \text{if } a = 1 \\ a, & \text{if } b = 1 \\ 0, & \text{otherwise,} \end{cases} = \begin{cases} \min(a, b), & \text{if } a = 1 \text{ or } b = 1 \\ 0, & \text{otherwise,} \end{cases}$$

is obviously a t-norm such that $Z \leqslant T$ for all t-norm T. Consequently,

$$\begin{aligned} Z^*(a, b) = 1 - Z(1 - a, 1 - b) &= \begin{cases} b, & \text{if } a = 0 \\ a, & \text{if } b = 0 \\ 1, & \text{otherwise,} \end{cases} \\ &= \begin{cases} \max(a, b), & \text{if } a = 0 \text{ or } b = 0 \\ 1, & \text{otherwise,} \end{cases} \end{aligned}$$

is a t-conorm such that $S \leqslant Z^*$ for all t-conorm S. Hence, for all t-norm T and all t-conorm S,

$$Z \leqslant T \leqslant \min \leqslant \max \leqslant S \leqslant Z^*.$$

[1] Notice that $T \leqslant S$ mean $T(a, b) \leqslant S(a, b)$, for all $(a, b) \in [0, 1] \times [0, 1]$.

Notice that min(max) is a continuous t-norm (t-conorm), but $Z(Z^*)$ is a discontinuous t-norm (t-conorm). The operations in [0, 1] given by

- $T_{prod}(a, b) = \ prod\,(a, b) = a\,\cdot\,b$
- $T_W(a, b) = W(a, b) = \max(0, a + b - 1) = (\max(0, Sum - 1))(a, b)$,

are also continuous t-norms. Then, the dual operations,

- $T^*_{prod}(a, b) = 1 - T_{prod}(1-a, 1-b) = 1-(1-a)\cdot(1-b) = a+b-a\cdot b = (Sum - prod)(a, b)$
- $W^*(a, b) = 1 - W(1 - a, 1 - b) = \min(1, a + b) = \min(1,\ Sum(a, b))$

are continuous t-conorms. Since it is easy to prove that

$$W \leqslant T_{prod} \leqslant \min,$$

it follows $\max \leqslant T^*_{prod} \leqslant W^*$, and

$$Z \leqslant W \leqslant T_{prod} \leqslant \min \leqslant \max \leqslant T^*_{prod} \leqslant W^*.$$

Remark 2.2.21 Since $Z(0.5, 0.5) = 0$, t-norm Z has zero-divisors. Analogously, from

$$W(a, b) = 0 \Leftrightarrow \max(0, a + b - 1) = 0 \Leftrightarrow a + b \leqslant 1,$$

it follows that t-norm W has zero-divisors, for example, $W(0.5, 0.4) = 0$; t-norms min and T_{prod} do not have zero-divisors:

- $T_{prod}(a, b) = 0 \Leftrightarrow a = 0$, or $b = 0$
- $\min(a, b) = 0 \Leftrightarrow a = 0$, or $b = 0$

Proposition 2.2.22 *The only idempotent t-norm, i.e.* $T(a, a) = a$*, for all* $a \in [0, 1]$*, is* $T = \min$.

Proof If T is idempotent, $\min(a, b) = T(\min(a, b), \min(a, b)) \leqslant T(a, b)$ since $\min(a, b) \leqslant a$, and $\min(a, b) \leqslant b$. Hence, $\min(a, b) \leqslant T(a, b) \leqslant \min(a, b)$ implies $T = \min$. □

Proposition 2.2.23 *The only idempotent t-conorm, i.e.* $S(a, a) = a$*, for all* $a \in [0, 1]$*, is* $S = \max$.

Proof If S is idempotent, $\max(a, b) = S(\max(a, b), \max(a, b)) \geq S(a, b)$, since $\max(a, b) \geq a$, and $\max(a, b) \geq b$. Hence, $\max(a, b) \geq S(a, b) \geq \max(a, b)$ implies $S = \max$. □

Remark 2.2.24 t-norms can be continuous, like min, T_{prod}, and W, or discontinuous, like Z. They can have zero-divisors, like W and Z, or not like min and T_{prod}. They can have all elements in [0, 1] idempotent (only $T = \min$), only have the idempotents

0 and 1 (like T_{prod} and W), or have some idempotents different from 0 and 1. In any case, since it is always $T(0, 0) = 0$ and $T(1, 1) = 1$, 0 and 1 are idempotent elements for all t-norms.

Remark 2.2.25 Analogous considerations can be made for t-conorms. There are discontinuous t-conorms like Z^*, and continuous ones like T^*_{prod} and W^*. The only for which all elements in [0, 1] are idempotent is $S = \max$. Since $S(0, 0) = 0$ and $S(1, 1) = 1$, 0 and 1 are always idempotent, and there are t-conorms that only have these two idempotents (like T^*_{prod} and W^*), as well as those that have some idempotents different from 0,1. There are t-conorms without one-divisors, like max and T^*_{prod}, and t-conorms with one-divisors like W^*, for example, $W^*(0.5, 0.5) = \min(1, 1) = 1$.

Remark 2.2.26 There is not a characterization theorem for all t-norms (t-conorms), *but it is a characterization of the continuous t-norms* (t-conorms) that will be presented by means of the following, and easy to prove, results:

- If $\varphi : [0, 1] \to [0, 1]$ verifies, (1) If $x \leqslant y$, then $\varphi(x) \leqslant \varphi(y)$, (2) φ is bijective, (3) $\varphi(0) = 0, \varphi(1) = 1$ (φ is an order-automorphism of the ordered interval $([0, 1], \leqslant)$), and T is a t-norm, then $T_\varphi = \varphi^{-1} \circ T \circ (\varphi \times \varphi)$ is also a t-norm. Given T, the set $\{T_\varphi; \varphi$ an order-automorphism$\}$ is called the *family* of T.
- T is a continuous t-norm if and only if all t-norms T_φ are continuous.
- If S is a t-conorm, then $S_\varphi = \varphi^{-1} \circ S \circ (\varphi \times \varphi)$ is also a t-conorm, and S is continuous if and only if all t-conorms S_φ are continuous, the set $\{S_\varphi; \varphi$ an order-automorphism$\}$ is called the *family* of S.

In particular,

- The family of $T = \min$, is reduced to the only t-norm min, since $\varphi^{-1}(\min(\varphi(a), \varphi(b)) = \min(\varphi^{-1}(\varphi(a)), \varphi^{-1}(\varphi(b))) = \min(a, b)$
- The family of T_{prod} contains all continuous t-norms of the form $\text{prod}_\varphi(a, b) = \varphi^{-1}(\varphi(a).(\varphi(b))$.
- The family of W contains all t-norms of the form $W_\varphi(a, b) = \varphi^{-1}(W(\varphi(a), \varphi(b)) = \varphi^{-1}(\max(0, \varphi(a) + \varphi(b) - 1))$, and all of them are continuous t-norms. Notice that no t-norm in the family $\{prod_\varphi\}$ has zero-divisors, since $prod_\varphi(a, b) = 0 \Leftrightarrow \varphi(a) \cdot \varphi(b) = 0 \Leftrightarrow \varphi(a) = 0$ or $\varphi(b) = 0 \Leftrightarrow a = 0$, or $b = 0$. Instead all t-norms W_φ have zero-divisors, since $W_\varphi(a, b) = 0 \Leftrightarrow \max(0, \varphi(a) + \varphi(b) - 1) \Leftrightarrow \varphi(a) + \varphi(b) \leqslant 1$. Of course, neither t-norms $prod_\varphi$, nor W_φ, have more idempotents than 0 and 1:
- $a = W_\varphi(a, a) = \varphi^{-1}(\max(0, 2\varphi(a) - 1)) \Leftrightarrow \varphi(a) = \max((0, 2\varphi(a) - 1) \Leftrightarrow \varphi(a) = 0$ or $\varphi(a) = 1$ or $a = 0$ or $a = 1$.
- $a = prod_\varphi(a, a) = \varphi^{-1}(\varphi(a).(\varphi(a)) \Leftrightarrow \varphi(a) = \varphi(a).(\varphi(a) \Leftrightarrow \varphi(a) = 0$ or $\varphi(a) = 1$ or $a = 0$ or $a = 1$.

Analogously,

- The family of $S = \max$, only contains this t-conorm.
- The family of T^*_{prod} contains all t-conorms of the form

$$prod^*_\varphi(a, a) = \varphi^{-1}(prod^*(\varphi(a), \varphi(b))) = \varphi^{-1}(\varphi(a) + \varphi(b) - \varphi(a) \cdot \varphi(b))$$

- The family of W^* contains all t-conorms of the form

$$W^*(a, a) = \varphi^{-1}(W^*(\varphi(a), \varphi(b)) = \varphi^{-1}(\min(1, \varphi(a) + \varphi(b)))$$

Remark 2.2.27 The order-automorphism φ plays the role of a functional parameter. By taking, $\varphi(x) = x^r$, it follows, for example,

$$W_\varphi(a, b) = \sqrt[r]{\max((0, a^r + b^r - 1)}, \quad W^*_\varphi(a, b) = \sqrt[r]{\min((1, a^r + b^r)}$$

giving a family of t-norms (t-conorms) depending on the numerical parameter $r > 0$. Notice that with $\varphi(x) = x^r$,

$$Prod_\varphi(a, b) = \sqrt[r]{a^r \cdot b^r} = a \cdot b = Prod(a, b),$$

but

$$Prod^*_\varphi(a, b) = \varphi^{-1}(\varphi(a) + \varphi(b) - \varphi(a) \cdot \varphi(b)) = \sqrt[r]{a^r + b^r - a^r \cdot b^r}.$$

2.2.6 Strong Negations

As it was said before, an strong negation is a function $N : [0, 1] \to [0, 1]$ such that

- $N(0) = 1$
- If $a \leqslant b$, then $N(b) \leqslant N(a)$
- $N(N(a)) = a$, for all $a \in [0, 1]$, or $N^2 = \text{id}$.

Notice that $N^2 = \text{id}$ is equivalent to $N = N^{-1}$, that shows N is a continuous function: It is $N(1) = N(N(0)) = 0$, and if $a < b$ it should be $N(b) < N(a)$ since $N(b) = N(a)$ would imply $N(N(b)) = N(N(a))$, or $a = b$. Hence, N is strictly decreasing.

Since N is continuous, the equation $N(x) = x$ has solutions, but there is only one. Suppose $N(x_1) = x_1$ and $N(x_2) = x_2$. Either $x_1 \leqslant x_2$, or $x_2 < x_1$. In the first case, it follows $N(x_2) \leqslant N(x_1)$, or $x_2 \leqslant x_1$, and $x_1 = x_2$. In the second case, $N(x_1) < N(x_2)$, or $x_1 < x_2$, that is absurd. Then, each strong negation has a single fixed point $N(n) = n$, in the open interval $(0, 1)$, since $N(0) = 1$, $N(1) = 0$ show that 0 and 1 are not fixed points.

Remark 2.2.28 In the classical case (a Boolean algebra L, or a power set $\mathbb{P}(X)$), the transformation

$$F : \mathbb{P}(X) \to \mathbb{P}(X),\ F(A) = A^c,$$

has no fixed points, since $A = A^c$ implies $A \cap A = A \cap A^c$, or $A = \varnothing$, and $\varnothing^c = X$. Nevertheless, with fuzzy sets

$$F : [0, 1]^X \to [0, 1]^X,\ F(\mu) = \mu' = N \circ \mu,$$

the equation $\mu = \mu'$, $N(\mu(x)) = \mu(x)$ for all x in X, has the only solution $\mu(x) = n$ for all x in X, that is $\mu = \mu_n$, with $n = N(n)$ the fix point of N.

In the fuzzy case, that mapping F shows a kind of symmetry that is not in the crisp case.

An order-automorphism of the ordered unit interval $([0, 1], \leqslant)$ $\varphi : [0, 1] \to [0, 1]$, verifies by definition,

- $\varphi(0) = 0, \varphi(1) = 1$
- If $a < b$, then $\varphi(a) < \varphi(b)$.

Hence, φ is continuous, and its inverse function φ^{-1} verifies,

- $\varphi^{-1}(0) = 0, \varphi^{-1}(1) = 1$
- If $a < b$, then $\varphi^{-1}(a) < \varphi^{-1}(b)$.

Let us denote by N_φ the function $N_\varphi : [0, 1] \to [0, 1]$ defined by

$$N_\varphi(a) = \varphi^{-1}(1 - \varphi(a)), \quad \text{for all } a \in [0, 1]$$

Proposition 2.2.29 *N_φ is a strong negation.*

Proof It is $N_\varphi(0) = \varphi^{-1}(1 - \varphi(0)) = \varphi^{-1}(1) = 1$. In addition, if $a \leqslant b$, it follows $1 - \varphi(b) \leqslant 1 - \varphi(a)$, and $\varphi^{-1}(1 - \varphi(b)) \leqslant \varphi^{-1}(1 - \varphi(a))$, or $N_\varphi(b) \leqslant N_\varphi(a)$. Finally,

$$\begin{aligned} N_\varphi(N_\varphi(a)) &= N_\varphi(\varphi^{-1}(1 - \varphi(a))) = \varphi^{-1}(1 - \varphi(\varphi^{-1}(1 - \varphi(a)))) \\ &= \varphi^{-1}(1 - 1 + \varphi(a)) = \varphi^{-1}(\varphi(a)) = a \end{aligned}$$

□

Theorem 2.2.30 *If N is a strong negation, there exist order-automorphisms φ such that $N = N_\varphi$.*

Proof Let it be $n = N(n) \in (0, 1)$ the fixed point of N, and consider an strictly non-decreasing function $h : [0, n] \to [0, \frac{1}{2}]$ such that $h(0) = 0$ and $h(n) = \frac{1}{2}$. With h define the function $\varphi : [0, 1] \to [0, 1]$ by

$$\varphi(x) = \begin{cases} h(x), & \text{if } x \in [0, n] \\ 1 - h(N(x)), & \text{if } x \in (n, 1]. \end{cases}$$

This function φ is, obviously, continuous, strictly increasing,[2] and verifies $\varphi(0) = h(0) = 0$, $\varphi(1) = 1 - h(N(1)) = 1 - h(0) = 1$. Then

- If $x \in [0, n)$, or $N(x) \in (n, 1]$, $\varphi(N(x)) = 1 - h(x) = 1 - \varphi(x)$, and $N(x) = \varphi^{-1}(1 - \varphi(x))$.
- If $x = n$, $N(n) = n = h^{-1}(\frac{1}{2}) = \varphi^{-1}(\frac{1}{2})$, or $N(n) = \varphi^{-1}(1 - \varphi(x)$
- If $x \in (n, 1]$, or $N(x) \in [0, n)$, $\varphi(N(x)) + \varphi(x) = h(N(x)) + 1 - h(N(x)) = 1$

In conclusion, $N(x) = \varphi^{-1}(1 - \varphi(x))$, for all x in $[0, 1]$, or $N = N_\varphi$. $\square$

Notice that the proof of last theorem shows clearly that the order-automorphism φ such that $N = N_\varphi$ is not unique. Notice also that with $\varphi = \mathrm{id}_{[0,1]}$ it follows $N(x) = 1 - x$, the fundamental strong negation, with which it results $N = N_\varphi = \varphi^{-1} \circ (1 - \mathrm{id}_{[0,1]}) \circ \varphi = \varphi^{-1} \circ N \circ \varphi$, that is, all strong negations belong to the family of $N_0(x) = 1 - x$. Nevertheless, in all cases it is $n = \varphi^{-1}(\frac{1}{2})$ the fixed point of N_φ.

If $\varphi(x) = x^2$, it results $N_\varphi(x) = \sqrt{1 - x^2}$, called the circular negation. If $\varphi(x) = \frac{2x}{1+x}$, or $\varphi^{-1}(x) = \frac{x}{2-x}$, it follows $N_\varphi(x) = \varphi^{-1}(1 - \frac{2x}{1+x}) = \varphi^{-1}(\frac{1-x}{1+x}) = \frac{1-x}{1+3x}$, that is the strong negation N_3 of the before mentioned Sugeno's negations.

With $\varphi(x) = \frac{1}{\lambda} \ln(1 + \lambda x^\alpha)$, $\lambda > -1$, $\alpha > 0$, it follows the bi-parametric family $N_\varphi(x) = (\frac{1-x^\alpha}{1+\lambda x^\alpha})^{\frac{1}{\alpha}}$, where with $\alpha = 1$ it is obtained the Sugeno's family of strong negations $N_\lambda = \frac{1-x}{1+\lambda x} (\lambda > -1)$ that only depends on one single parameter.

Remark 2.2.31 The only linear strong negation N is $N = N_0$, since from $N(a) = \alpha a + \beta$, with $N(0) = 1 = \beta$ and $N(1) = 0 = \alpha + 1$, follows $\alpha = -1$ and $N(a) = 1 - a$.

Remark 2.2.32 The only "rational" strong negations N of the form $N(x) = \frac{ax+b}{cx+d}$, a, b different of 0, are those $N_\lambda (\lambda > -1)$ in the Sugeno's family. It follows from:

- $N(0) = 1 = \frac{b}{d}$, or $d = b$
- $N(1) = 0 = \frac{a+b}{c+d}$, or $a = -b$

that gives

$$N(x) = \frac{-bx + b}{cx + b} = \frac{b(1 - x)}{cx + b} = \frac{1 - x}{1 + \frac{c}{b} x}.$$

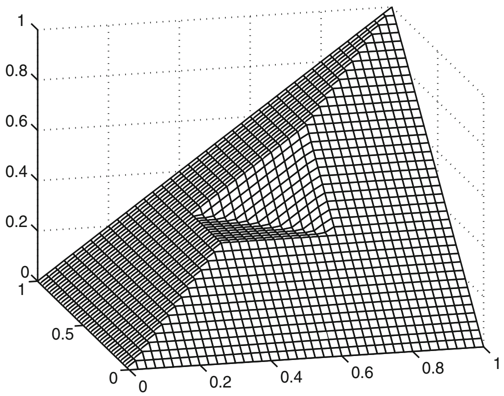
To have $0 \leqslant N(x) \leqslant 1$, it should be $1 - x \leqslant 1 + \frac{c}{b}$. But $-1 = \frac{c}{b}$ implies $N(x) = 1$, that is not an strong negation. Hence it is $-1 < \frac{c}{b}$, and with $\lambda = \frac{c}{b}$, it follows $N(x) = \frac{1-x}{1+\lambda x} = N_\lambda(x)$, with $-1 < \lambda$.

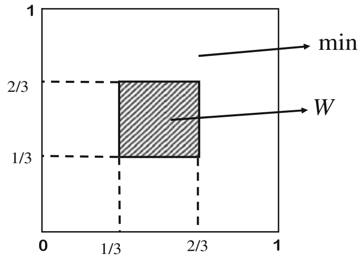
[2] For $(x < y)$ is evident that $\varphi(x) < \varphi(y)$ if either $x, y \in [0, n]$, or $x, y \in (n, 1]$. If $x \in [0, n]$, $y \in (n, 1]$ and $x < y$, since $h(x) + h(N(x)) < 1$, it is $h(x) < 1 - h(N(x))$, or $\varphi(x) < \varphi(y)$.

2.2.7 Continuous T-Norms and T-Conorms

As it was said, the only t-norm that is idempotent for all a in [0, 1], is $T = \min$, and the t-norms in $\{prod\} \cup \{W\}$ only have the idempotents 0 and 1. As it was also said, there are t-norms with several (but not all) idempotent elements. For example, the function

$$T(x, y) = \begin{cases} \frac{1}{3} + \frac{1}{3} W(3x - 1, 3y - 1), & \text{if } (x, y) \in [\frac{1}{3}, \frac{2}{3}]^2 \\ \min(x, y), & \text{otherwise ,} \end{cases}$$





that as it is easy to prove is a t-norm, verifies

- $T(x, x) = \min(x, x) = x$, if $x \notin [\frac{1}{3}, \frac{2}{3}]$
- $T(\frac{1}{3}, \frac{1}{3}) = \frac{1}{3} + \frac{1}{3} W(0, 0) = \frac{1}{3} + \frac{1}{3}.0 = \frac{1}{3}$
- $T(\frac{2}{3}, \frac{2}{3}) = \frac{1}{3} + \frac{1}{3} W(1, 1) = \frac{1}{3} + \frac{1}{3} = \frac{2}{3}$
- $T(\frac{1}{2}, \frac{1}{2}) = \frac{1}{3} + \frac{1}{3} W(\frac{3}{2} - 1, \frac{3}{2} - 1) = \frac{1}{3} + \frac{1}{3} W(\frac{1}{2}, \frac{1}{2}) = \frac{1}{3} + \frac{1}{3}.0 = \frac{1}{3} \neq \frac{1}{2}$
- etc.

that is, all elements in $[0, 1] - [\frac{1}{3}, \frac{2}{3}]$, as well $\frac{1}{3}$ and $\frac{2}{3}$ are idempotent for T, and the elements in $(\frac{1}{3}, \frac{2}{3})$ are not idempotent.

Look that an analogous result is obtained when changing W by prod in the above expression of T. Without proof it follows the theorem that completely characterizes all continuous t-norms.

Theorem 2.2.33 *T is a continuous t-norm if and only if,*

1. $T = \min$, *T is in the family of min*
2. $T = prod_\varphi$, *T is in the family of prod*
3. $T = W_\varphi$, *T is in the family of W*
4. *There exist an index set (finite or countable infinite), a family of pairwise disjoint open intervals in* [0, 1], $\{(a_1, b_1); i \in I\}$, *and a family of t-norms* $T_i \in \{prod\} \cup \{W\} (i \in I)$, *such that*

$$T(x, y) = \begin{cases} a_i + (b_i - a_i)T_i(\frac{x-a_i}{b_i-a_i}, \frac{y-a_i}{b_i-a_i}), & \text{if } (x, y) \in [a_i, b_i]^2 \\ \min(x, y), & \text{otherwise,} \end{cases}$$

for any x, y *in* $[0, 1$.

The continuous t-norm of the fourth type are called *ordinal-sums* of the continuous t-norms $T_i \in \{prod\} \cup \{W\}$.

Remark 2.2.34 Why the names t-norm and t-conorm? The "t" comes from "triangular", because these functions were introduced to formalize the triangular property of probabilistic distances, i.e. distances whose values are something like the probability that the numerical distance between two points is less than a given number. They were introduced by Karl Menger with the name triangular-norms without considering associativity.

The name of t-conorm refers to the duality with a t-norm, since S is a t-conorm if and only if $1 - S(1 - x, 1 - y)$ is a t-norm. In general, it should be pointed out that, for each strong negation N, S is a t-conorm if and only if $N \circ S \circ (N \times N)$ is a t-norm.

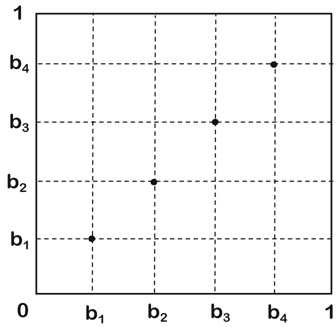
Theorem 2.2.35 S *is a continuous t-conorm if and only if,*

1. $S = \max$, *S is in the family of max*
2. $S = prod^*_\varphi$, *S is in the family of prod**
3. $S = W^*_\varphi$, *S is in the family of* W^*
4. *There exist an index set (finite or countable infinite), a family of pairwise disjoint open intervals of* $[0, 1], \{(a_1, b_1); i \in I\}$, *and a family of t-conorms* $S_i \in \{prod^*\} \cup \{W^*\}(i \in I)$, *such that*

$$S(x, y) = \begin{cases} a_i + (b_i - a_i)S_i(\frac{x-a_i}{b_i-a_i}, \frac{y-a_i}{b_i-a_i}), & \text{if } (x, y) \in [a_i, b_i]^2 \\ \max(x, y), & \text{otherwise,} \end{cases}$$

for any x, y *in* $[0, 1]$.

Remark 2.2.36 Notice that with both ordinal-sums of t-norms and of t-conorms, provided it is $[0, 1] = [0, b_1] \cup [b_1, b_2] \cdots \cup [b_{n-1}, b_n] \cup [b_n, 1]$, a finite partition of the unit interval $[0, 1]$, like, for example



the only idempotent elements are b_1, b_2, b_3, b_4, etc., as well as 0 and 1, that is, the points giving the partition of [0, 1].

Remark 2.2.37 Although currently only continuous t-norms in $\{\min\} \cup \{prod\} \cup \{W\}$ are taken into account in both theoretic fuzzy logic and its applications, it should be pointed out that provided there are, at least, two statements 'x is P' and 'x is Q' such that $\mu_P(x), \mu_Q(x) \notin \{0, 1\}$ and $\mu_{P\ and\ P} = \mu_P$, $\mu_{Q\ and\ Q} = \mu_Q$, the only possibility for representing $\mu \cdot \sigma = T \circ (\mu \times \sigma)$, is by taking as continuous t-norm T an ordinal-sum with the single interval $(\min(\mu_P(x), \mu_Q(x)), \max(\mu_P(x), \mu_Q(x)))$.

Remark 2.2.38 Which t-norms are strictly non-decreasing in the sense that if $0 < a < b < 1$, then $T(a, c) < T(b, c)$ for all $c \in [0, 1]$?

- If $T = \min$, the answer is negative. For example, $0.3 < 0.5$, but $\min(0.2, 0.3) = \min(0.2, 0.5) = 0.2$
- If $T = W_\varphi$, the answer is also negative. For example,$0.3 < 0.5$, but $W(0.2, 0.3) = W(0.2, 0.5) = 0$
- If $T = prod_\varphi$, the answer is positive, since: $a < b \Rightarrow \varphi(a) < \varphi(b) \Rightarrow \varphi(a) \cdot \varphi(c) < \varphi(b) \cdot \varphi(c) \Rightarrow \varphi^{-1}(\varphi(a) \cdot \varphi(c))) < \varphi^{-1}(\varphi(b) \cdot \varphi(c)))$, or $prod_\varphi(a, c) < prod_\varphi(b, c)$, because $\varphi(c) \in (0, 1]$.
- If T is an ordinal-sum, it can't be strictly non-decreasing because of the values it takes with min.

  Analogously, the only t-conorms that are strictly non-decreasing are those in $\{prod^*_\varphi\}$.

2.2.8 Laws of Fuzzy Sets

As it was said, in all standard algebras $([0, 1]^X, T, S, N)$ of fuzzy sets the triplet (T, S, N) share the following common properties:

1. T and S are commutative and associative
2. 1 is neutral for T, and 0 is neutral for S
3. 0 is absorbent for T, and 1 is absorbent for S
4. For all T and S, it is $T \leqslant \min < \max \leqslant S$
5. Each T (S) is non decreasing in the two variables
6. N is involutive, strictly decreasing and such that $N(0) = 1$,

a list of properties that gives some basic laws for fuzzy sets in the standard algebras, like

- $\mu \cdot \sigma = \sigma \cdot \mu$, $\mu + \sigma = \sigma + \mu$,
- $\mu + (\sigma + \lambda) = (\mu + \sigma) + \lambda = (\sigma + \mu) + \lambda = \lambda + (\sigma + \mu)$
- $\mu \cdot \mu_1 = \mu$, $\mu + \mu_0 = \mu$, $\mu + \mu_1 = \mu_1$, $\mu \cdot \mu_0 = \mu_0$
- If $\mu \leqslant \sigma$, then $\mu \cdot \lambda \leqslant \sigma \cdot \lambda$, and $\lambda + \mu \leqslant \sigma + \lambda$
- etc.

Anyway, *a lot of laws typical of classical sets are not always valid* in all standard algebras of fuzzy sets. For example, $(\mathbb{P}(X), \cap, \cup, {}^c)$ is a Boolean algebra and no one $([0, 1]^X, T, S, N)$ is a Boolean algebra. In particular, $(\mathbb{P}(X), \cap, \cup)$ is a lattice and the only standard algebra that is a lattice is that with $T = \min$ and $S = \max$. Let us study in which standard algebras some laws of crisp sets do hold.

2.2.8.1 Distributive Laws

With classical sets it always do hold the two distributive laws

1. $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
2. $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

and the question is for which triplets (T, S, N) do hold the corresponding laws with fuzzy sets

1. $\mu \cdot (\sigma + \lambda) = \mu \cdot \sigma + \mu \cdot \lambda$,
2. $\mu + (\sigma \cdot \lambda) = (\mu + \sigma) \cdot (\mu + \lambda)$.

This questions correspond to solve the functional equations in the unknowns T and S:

$$T(a, S(b, c)) = S(T(a, b), T(a, c)) \tag{2.3}$$

$$S(a, T(b, c)) = T(S(a, b), S(a, c)) \tag{2.4}$$

for all a, b, c in $[0, 1]$.

Lemma 2.2.39 *Equation* (2.3) *does hold if and only if* $S = \max$.

Proof With $b = c = 1$, is $T(a, S(1, 1)) = S(T(a, 1), T(a, 1))$ or $a = S(a, a)$. That is $S = \max$.

Provided $S = \max$, the equation is $T(a, \max(b, c)) = \max(T(a, b), T(a, c))$. If either $b \leqslant c$ or $c \leqslant b$, it is immediate to check its validity for all t-norm T. □

Lemma 2.2.40 *Equation* (2.4) *does hold if and only if* $T = \min$.

Proof With $b = c = 0$, is $S(a, 0) = a = T(a, a)$. That is $T = \min$.

Provided $T = \min$, the equation is $S(a, \min(b, c)) = \min(S(a, b), S(a, c))$. If either $b \leqslant c$ or $c \leqslant b$, it is immediate to check its validity for all t-conorm T. □

Hence,

- In all standard algebras with $(T, \max)$, it holds $\mu \cdot (\sigma + \lambda) = \mu \cdot \sigma + \mu \cdot \lambda$,
- In all standard algebras with $(\min, S)$, it holds $\mu + (\sigma \cdot \lambda) = (\mu + \sigma) \cdot (\mu + \lambda)$
- The two distributive laws (2.3) and (2.4) do jointly hold if and only if $T = \min$ and $S = \max$.

2.2.8.2 De Morgan Laws

With classical sets it always do hold the De Morgan, or duality, laws

1. $A \cup B = (A^c \cap B^c)^c$
2. $A \cap B = (A^c \cup B^c)^c$

showing that one of the two operators $\cap, \cup$ can be defined by the other and the complementation. With fuzzy sets in standard algebras, these laws are

1. $\mu \cdot \sigma = (\mu' + \sigma')'$, or $(\mu \cdot \sigma)' = \mu' + \sigma'$
2. $\mu + \sigma = (\mu' \cdot \sigma')'$, or $(\mu + \sigma)' = \mu' \cdot \sigma'$,

that correspond to the functional equations in the unknowns T, S and N

- $T(a, b) = N(S(N(a), N(b)))$
- $S(a, b) = N(T(N(a), T(b)))$

for all a, b in $[0, 1]$.

Obviously, law (1) does hold if and only if $T = N \circ S \circ (N \times N)$ and law (2) does hold if and only if $S = N \circ T \circ (N \times N)$, two formulas that are equivalent since, for example, from the first it follows (with $N^2 = \text{id}$) $N \circ T = S \circ (N \times N)$ or $N \circ S = T \circ (N \times N)$, that is, the second formula.

Hence, *the two* De Morgan *laws hold in an algebra given by the triplet* (T, S, N) *if and only if* $T = N \circ S \circ (N \times N)$ *, that is,* T *and* S *are* N*-dual.*

2.2.8.3 Restricted Non-contradiction Principle $\mu \cdot \mu' = \mu_0$

With classical sets it always holds $A \cap A^c = \emptyset$. With fuzzy sets, when is it

$$\mu \cdot \mu' = \mu_0 \ ?$$

The equation to be solved is

$$T(a, N(a)) = 0, \quad \text{for all a } \in [0, 1]$$

in the unknowns T and N.

Theorem 2.2.41 *If* T *is a continuous t-norm, and* N *is a strong negation, it is* $T(a, N(a)) = 0$ *for all* $a \in [0, 1]$*, if and only if* $T = W_\varphi$ *and* $N \leqslant N_\varphi$*.*

Proof With the fixed point $n \in (0, 1)$ of N, it follows $T(n, n) = 0$, that is, T has zero-divisors. Hence, $T = W_\varphi$, and $W_\varphi(a, N(a)) = \varphi^{-1}(\max(0, \varphi(a)+\varphi(N(a))-1)) = 0$, or $\max(0, \varphi(a) + \varphi(N(a)) - 1) = 0$, or $\varphi(a) + \varphi(N(a)) - 1 \leqslant 0$, that implies $\varphi(N(a)) \leqslant 1 - \varphi(a)$, or $N(a) \leqslant \varphi^{-1}(1 - \varphi(a)) = N_\varphi(a)$, for all $a \in [0, 1]$. Hence $N \leqslant N_\varphi$. The reciprocal is a simple calculation. $\square$

Then, the (restricted) non-contradiction principle $\mu \cdot \mu' = \mu_0$ holds if and only if $T = W_\varphi$ and $N \leqslant N_\varphi$, for any order-automorphism φ and any t-conorm S. For example, it holds (with $\varphi = \text{id}$) if $T = W$, $S = \max$, $N = N_0$, and it does not hold provided $T = \min$, or $T = prod_\varphi$.

2.2.8.4 Restricted Excluded-Middle Principle $\mu + \mu' = \mu_1$

With classical sets it always holds $A \cup A^c = X$. When is it $\mu + \mu' = \mu_1$ for fuzzy sets? When it does hold the equation $S(a, N(a)) = 1$ for all $a \in [0, 1]$?

Theorem 2.2.42 *If S is a continuous t-conorm, and N is a strong negation, it is $S(a, N(a)) = 1$ for all $a \in [0, 1]$, if and only if $S = W_\psi^*$ and $N_\psi \leqslant N$.*

Proof With $N(n) = n \in (0, 1)$, it follows $S(n, n) = 1$. That is $S = W_\psi^*$, and $1 = W_\psi^*(a, n(a)) = \psi^{-1}(\min(1, \psi(a) + \psi(N(a))))$, or $1 = \min(1, \psi(a) + \psi(N(a)))$. Hence, $1 \leqslant \psi(a)+\psi(N(a))$, or $N_\psi(a) = \psi^{-1}(1-\psi(a) \leqslant N(a))$. That is, $N_\psi \leqslant N$. The reciprocal is a simple calculation. □

Then, the (restricted) excluded-middle principle $\mu + \mu' = \mu_1$ holds if and only if $S = W_\psi^*$ and $N_\psi \leqslant N$, for any order automorphism ψ and any t-norm T. For example, it holds (with $\psi = \text{id}$) if $S = W^*$, $T = \min$, $N = N_0$, but it does not hold provided $S = \max$ or $S = prod^*$.

2.2.8.5 Both Restricted Principles of Non-contradiction and Excluded-Middle

From last theorems it immediately follows that,

Theorem 2.2.43 *In a standard algebra of fuzzy sets with a triplet (T, S, N), it holds $\mu \cdot \mu' = \mu_0$ and $\mu + \mu' = \mu_1$ if and only if $T = W_\varphi$, $S = W_\psi^*$, and $N_\psi \leqslant N \leqslant N_\varphi$.*

In particular, they hold if $T = W$, $S = W^*$, and $N = N_0$, or with $\varphi(x) = x^2$ and $\psi(x) = x$, they hold with the triplet:

$$T(x, y) = \sqrt{\max(0, x^2 + y^2 - 1)},\ S(x, y) = \min(1, x + y),$$

$$1 - x \leqslant N(x) \leqslant \sqrt{1 - x^2},$$

for all x, y in $[0, 1]$.

Of course, with $\varphi = \psi$, the principles hold with W_φ, W_φ^*, and $N = N_\varphi$.

2.2.8.6 Laws of Absorption

With classical sets, the absorption laws $A \cap (A \cup B) = A$, and $A \cup (A \cap B) = A$, always hold. With fuzzy sets in standard algebras, the formulas and respective equations

- $\mu \cdot (\mu + \sigma) = \mu$, $T(a, S(a, b)) = a$
- $\mu + (\mu \cdot \sigma) = \mu$, $S(a, T(a, b)) = a$

must be studied to find for which algebras these laws do hold.

Lemma 2.2.44 *If T and S are, respectively, a t-norm and a t-conorm, it is $T(a, S(a, b)) = a$ for all a, b in* $[0, 1]$ *if and only if* $T = \min$.

Proof If $T = \min$, since $a \leqslant S(a, b)$, it follows $\min(a, S(a, b)) = a$. With $b = 0$, the equation gives $T(a, a) = a$, and $= \min$. $\square$

Lemma 2.2.45 *If T and S are, respectively, a t-norm and a continuous t-conorm, it is $S(a, T(a, b)) = a$ for all a, b in* $[0, 1]$ *if and only if* $S = \max$.

Proof Since $T(a, b) \leqslant a$, it follows $\max(a, T(a, b)) = a$. With $b = 1$, the equation gives $S(a, a) = a$, and $= \max$. $\square$

Hence,

- The law $\mu \cdot (\mu + \sigma) = \mu$, holds for all S and $T = \min$
- The law $\mu + (\mu \cdot \sigma) = \mu$, holds for all T and $S = \max$
- The two laws hold jointly if and only if $T = \min$ and $S = \max$.

2.2.8.7 The Law of von Neumann

With classical sets it always holds the law of von Neumann, or law of the perfect repartition,

$$A = (A \cap B) \cup (A \cap B^c),$$

that follows from $A = A \cap X = A \cap (B \cup B^c) = (A \cap B) \cup (A \cap B^c)$, and generalizes that of the excluded-middle since $A = X$ implies $X = (B \cap X) \cup (X \cap B^c) = B \cup B^c$.

From that law, by duality it follows $A^c = (A \cap B)^c \cap (A \cap B^c)^c = (A^c \cup B^c) \cap (A^c \cup B)$, that is

$$A = (A \cup B) \cap (A \cup B^c),$$

a law that generalizes that of non-contradiction since $A = \varnothing$ implies $\varnothing = B \cap B^c$.

With fuzzy sets, the question is the validity of the laws

$$\mu = \mu \cdot \sigma + \mu \cdot \sigma', \quad \mu = (\mu + \sigma) \cdot (\mu + \sigma'),$$

or of the functional equations

$$a = S(T(a,b), T(a, N(b))), \quad a = T(S(a,b), S(a, N(b)))$$

Lemma 2.2.46 *The equation $a = S(T(a,b), T(a, N(b)))$ holds if and only if $T = prod_\varphi$, $S = W^*_\varphi$, $N = N_\varphi$.*

Proof If $T = prod_\varphi, S = W^*_\varphi, N = N_\varphi$, it is $S(T(a,b), T(a, N(b))) = W^*_\varphi(prod_\varphi(a,b), prod_\varphi(a, N_\varphi(b))) = \varphi^{-1}(W^*(\varphi(prod_\varphi(a,b)), \varphi(prod_\varphi(a, N_\varphi(b))))) = \varphi^{-1}(W^*(\varphi(a)\cdot\varphi(b), \varphi(a)\cdot\varphi(N_\varphi(b)))) = \varphi^{-1}(\min(1, \varphi(a)\cdot\varphi(b)+\varphi(a)\cdot\varphi(N_\varphi(b)))) = \varphi^{-1}(\min(1, \varphi(a)\cdot\varphi(b) + \varphi(a)\cdot(1-\varphi(b)) = \varphi^{-1}(\min(1, \varphi(a)\cdot\varphi(b) + \varphi(a) - \varphi(a)\cdot\varphi(b))) = \varphi^{-1}(\min(1, \varphi(a)) = a$.

The proof of the reciprocal will be avoided since it is technically complex. Let us say only that $a = S(T(a,b), T(a, N(b)))$ gives, with $a = 1$, $1 = S(b, N(b))$, that implies $S = W^*$ and $N_\varphi \leqslant N$. □

It can be also proven that $a = T(S(a,b), S(a, N(b)))$ if and only if $T = W_\varphi$, $S = prod^*_\varphi$, $N = N_\varphi$. Notice only that $a = 0$ gives $T(b, N(b)) = 0$, or $T = W_\varphi$ and $N = N_\varphi$.

Notice that the verification of von Neumann's law require, in the case of fuzzy sets, non-dual theories, like those given by the triplets $(prod_\varphi, W^*_\varphi, N_\varphi)$, and $(W_\varphi, prod^*_\varphi, N_\varphi)$.

2.2.8.8 Which Standard Algebra Is Closer to a Boolean Algebra?

The results in last section can be summarized in the following Table 2.1.

Hence, the algebras with the triplets (min, max, N) are the ones that preserve more structural Boolean properties. Indeed, these algebras preserve all the basic Boolean laws except those of non-contradiction and excluded-middle. They are distributive pseudo-complemented lattices, that is, De Morgan algebras that, in addition and like all algebras of fuzzy sets, verify the law of Kleene,

$$T(a, N(a)) \leqslant S(b, N(b)),$$

for all a, b in $[0, 1]$. The algebras given by the triplets (min, max, N) are De Morgan-Kleene algebras.

2.2.8.9 Last Comments

It can be considered, in addition to the structural Boolean laws, the cases that can be derived from them, for example,

$$(A \cap B^c)^c = B \cup (A^c \cap B^c),$$

Table 2.1 Basic boolean properties

	T	**S**	**N**
Lattice	Min	Max	All
Identity			
$T(a, 1) = a, T(a, 0) = 0$	All	–	–
$S(a, 0) = a, S(a, 1) = 1$	–	All	–
Commutativity			
$T(a, b) = T(b, a)$	All	–	–
$S(a, b) = S(b, a)$	–	All	–
Associativity			
$T(a, T(b, c)) = T(T(a, b), c)$	All	–	–
$S(a, S(b, c)) = S(S(a, b), c)$	–	All	–
Involution			
$N(N(a)) = a$	–	–	All
B1. Idempotency			
$T(a, a) = a$	Min	–	–
$S(a, a) = a$	–	Max	–
Distributivity			
$T(a, S(b, c)) = S(T(a, b), T(a, c))$	All	Max	–
$S(a, T(b, c)) = T(S(a, b), S(a, c))$	Min	All	–
Absorption			
$T(a, S(a, b)) = a$	Min	All	–
$S(a, T(a, b)) = a$	All	Max	–
Non-contradiction			
$T(a, N(a)) = 0$	W_φ	–	$N \leqslant N_\varphi$
Excluded-middle			
$S(a, N(a)) = 1$	–	W_φ^*	$N \geq N_\varphi$
De Morgan's laws			
$N(T(a, b)) = S(N(a), N(b))$	$T = N \circ S \circ N \times N$		
$N(S(a, b)) = T(N(a), N(b))$	$T = N \circ S \circ N \times N$		

that follows from $B \cup (A^c \cap B^c) = (B \cup A^c) \cap (B \cup B^c) = (B \cup A^c) \cap X = B \cup A^c = (A \cap B^c)^c$. This law, in fuzzy set theory, is translated by

$$(\mu \cdot \sigma')' = \sigma + \mu' \cdot \sigma'$$

or

$$N(T(a, N(b))) = S(b, T(N(a), N(b))),$$

that holds with $T = prod_\varphi$, $S = W_\varphi^*$, $N = N_\varphi$.

Nevertheless, not all derived law has solution within the standard algebras of fuzzy sets, as it is the case with $(A \cup A) \cap (A \cap A^c) = \varnothing$ (or $A \cap \varnothing = \varnothing$), since for

$$\mu \cdot \mu + \mu \cdot \mu' = \mu_0, \quad [\star]$$

there are no triplets (T, S, N) for which it can hold $S(T(a, a), T(a, N(a))) = 0$ for all a in [0, 1].

In the same vein, there are some laws that have solutions when different t-norms, t-conorms and strong negations are considered. For example, $(\mu + \mu) \cdot (\mu \cdot \mu') = \mu_0$, that comes from $(A \cup A) \cap (A \cap A^c) = \varnothing$, translated in the form $T_1(S(a, a), T_2(a, N(a))) = 0$, has infinite solutions like, for example, with an strong negation N, such that $N \leqslant N_0$, $T_1 = \min$, $T_2 = W$ and any t-conorm S, since $\min(S(a, a), T_2(a, N(a)) = T_2(a, N(a)) = W(a, N(a)) = \max(0, a + N(a) - 1) = 0$, because of $T_2(a, N(a)) \leqslant a \leqslant S(a, a)$, and $N(a) \leqslant 1 - a$, or $a + N(a) - 1 \leqslant 0$.

Another case is given by the classical (derived) laws

$$A \cap (A^c \cup B) = A \cap B, \quad A \cup (A^c \cap B) = A \cup B,$$

and the corresponding 'possible' fuzzy laws

$$\mu \cdot (\mu' + \sigma) = \mu \cdot \sigma, \;\; \mu + (\mu \cdot \sigma) = \mu + \sigma,$$

which functional equations

$$T_1(a, S(N(a), b)) = T_2(a, b), \;\; S_1(a, T(N(a), b)) = S_2(a, b),$$

do not have solutions with $T_1 = T_2$ and $S_1 = S_2$, respectively, but that with $N = N_0$, W and W^* do verify

- $W(a, W^*(1 - a, b)) = \max(0, \min(a, b)) = \min(a, b)$
- $W^*(a, W(1 - a, b)) = \min(1, \max(a, b)) = \max(a, b)$

that is, they have the solutions $(T_1 = W, S = W^*, T_2 = \min)$ and $(S_1 = W^*, T = W, S_2 = \max)$, respectively. Thus, it is possible to consider more complex algebras of fuzzy sets by means of n-tuples of the type $(T_1, \ldots, T_m; S_1, \ldots, S_r; N_1, \ldots, N_p)$.

Notwithstanding, there are more derived laws than $[\star]$ that have no solutions neither in standard algebras, nor with different t-norms, t-conorms, or different strong negations. The fact that no standard algebra of fuzzy sets is a Boolean algebra, makes impossible to simultaneously deal in such algebras with all formulas that are valid with classical sets.

2.2.9 Examples

Example 2.2.47 In a scale between 10 and 50 °C, the label 'cold' referred to temperature, is graduated by

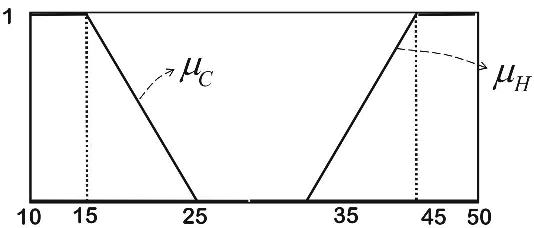
$$\mu_C(x) = \begin{cases} 1, & \text{if } 10 \leqslant x \leqslant 15 \\ \frac{25-x}{10}, & \text{if } 15 \leqslant x \leqslant 25 \\ 0, & \text{if } 25 \leqslant x \leqslant 50. \end{cases}$$

Which one of the following linguistic labels: *cold, hot, warm, more or less cold, more or less warm*, is the more adequate for the temperatures of 20, 21 and 22 °C?

Solution. With C = cold, it is $\mu_C(20) = \frac{5}{10} = 0.5$ and $\mu_{more\ or\ less\ C}(20) = \sqrt{0.5} = 0.71$. To obtain $H = hot$, we can compute μ_H as the opposite of μ_C:

$$\mu_H(x) = \mu_C(50 + 10 - x) = \mu_C(60 - x) = \begin{cases} 1, & \text{if } 45 \leqslant x \leqslant 50 \\ \frac{x-35}{10}, & \text{if } 35 \leqslant x \leqslant 45 \\ 0, & \text{if } 10 \leqslant x \leqslant 35 \end{cases}$$

graphically



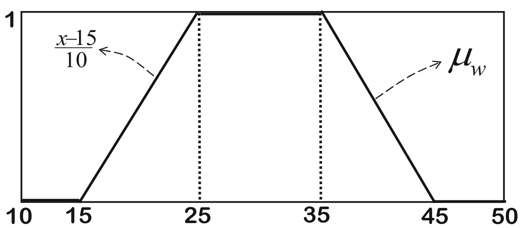
By defining, as it is usual, warm = not cold and not hot, that is

$$\mu_w = \mu'_{cold} \cdot \mu'_{hot}$$

with $\cdot$ represented by min, and $'$ by N_0,

$$\mu_w(x) = \min(1 - \mu_{cold}(x), 1 - \mu_{hot}(x)),$$

for all $x \in [10, 50]$, it results



Then:

- $x = 20$, gives $\mu_c(20) = \frac{5}{10} = 0.5$, $\mu_H(20) = 0$, $\mu_W(20) = \frac{5}{10} = 0.5$,
- $x = 21$, gives $\mu_c(21) = \frac{4}{10} = 0.4$, $\mu_H(21) = 0$, $\mu_W(21) = \frac{6}{10} = 0.6$,
- $x = 22$, gives $\mu_c(22) = \frac{3}{10} = 0.3$, $\mu_H(22) = 0$, $\mu_W(22) = \frac{7}{10} = 0.7$,

and

- $\mu_{more\ or\ less\ cold}(20) = \sqrt{\mu_c(20)} = 0.71$,
  $\mu_{more\ or\ less\ warm}(20) = \sqrt{\mu_W(20)} = 0.71$
- $\mu_{more\ or\ less\ ;cold}(21) = 0.63$, $\mu_{more\ or\ less\ ;warm}(21) = 0.77$,
- $\mu_{more\ or\ less\ ;cold}(22) = 0.55$, $\mu_{more\ or\ less\ ;warm}(22) = 0.84$.

Hence

- The more adequate linguistic label for $x = 20$, cannot be decided but it could be either 'not cold', or 'not warm'. Since, it is not hot at all, we can take 'not cold'.
- For $x = 21$, is 'mol warm' (mol = more or less)
- For $x = 22$, is 'mol warm'

Example 2.2.48 On the age of a person p, it is known that

$$37 \leqslant Age(p) \leqslant 41,$$

and neither $Age(p) \leqslant 32$, nor $43 \leqslant Age(p)$. What can be said on the degree up to which it could be $Age(p) = 35$, and $Age(p) = 42$?

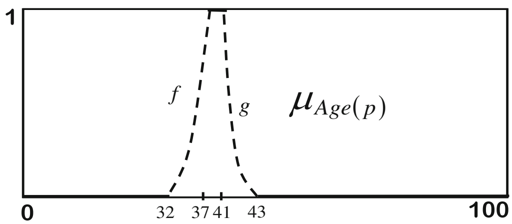
Solution. What is unknown is the variation of $Age(p)$ between 32 and 37, as well as between 41 and 43. Since Age varies continuously, we can suppose there are two functions

$$f : [32, 37] \to [0, 1], g : [41, 43] \to [0, 1]$$

such that $f(32) = 0, f(37) = 1, g(41) = 1, g(43) = 0$, with f strictly non-decreasing, and g strictly decreasing. Then, we can define $\mu_{Age(p)} : [0, 100] \to [0, 1]$, by

$$\mu_{Age(p)}(x) = \begin{cases} 0, & \text{if } x \in [0, 32] \cup [43, 100] \\ 1, & \text{if } x \in [37, 41] \\ f(x), & \text{if } x \in [32, 37] \\ g(x), & \text{if } x \in [41, 43] \end{cases}$$

and $\mu_{Age(p)}(35) = f(35)$, $\mu_{Age(p)}(42) = g(42)$. Graphically



To determine f and g more information is needed, but in the absence of it, we can decide to take the linear models $f(x) = \frac{x-32}{5}$, and $g(x) = \frac{43-x}{2}$, with which

$$\mu_{Age(p)}(35) = \frac{3}{5}, \quad \mu_{Age(p)}(42) = \frac{1}{2}.$$

As it will be seen later on, 0.6 is the *possibility* that $Age(p) = 35$, and 0.5 that of $Age(p) = 42$. Hence, it seems a little bit more possible that it be '$Age(p) = 42$' than '$Age(p) = 35$'.

Example 2.2.49 Knowing that Height(John) = 175 cm, and Height(Peter) = 180 cm, consider the two statements:

p = It is false that John is not very tall or is more or less short
q = It is false that Peter is not very tall or is more or less short.

which is more true?

Solution. Both statements can be written by

Is false that x is P,
with P = '(not very tall) or (more or less short)'.

Hence

$$\mu_P(x) = S(\mu'_{very\ tall}(x), \mu_{not\ short}(x)) = S(N(\mu_{tall}(x)^2), \sqrt{\mu_{tall}(A(x))}\ ),$$

with a continuous t-conorm S, a strong negation N, and a symmetry A on X, provided x varies in a scale of heights.

What should be compared are the two values $N(\mu_P(175))$ and $N(\mu_P(180))$, and for that it is needed to know μ_{tall}. Let us take

$$\mu_{tall}(x) = \begin{cases} 0, & \text{if } x \in [0, 150] \\ \text{strictly non decreasing}, & \text{if } x \in [150, 190] \\ 1, & \text{if } x \in [190, 210] \end{cases}$$

with, perhaps, $\mu_{tall}(x) = 0.025x - 3.75$, $x \in [150, 190]$, if we need to have numbers.

Hence, with $A(x) = 210 - x$, it is $A(175) = 210 - 175 = 35$, and $\mu_{tall}(35) = 0$, as well as $A(180) = 220 - 180 = 30$, and $\mu_{tall}(30) = 0$, because of that

$$\mu_P(175) = S(N(\mu_{tall}(175)), 0) = N(\mu_{tall}(175)^2)$$
$$\mu_P(180) = S(N(\mu_{tall}(180)), 0) = N(\mu_{tall}(180)^2).$$

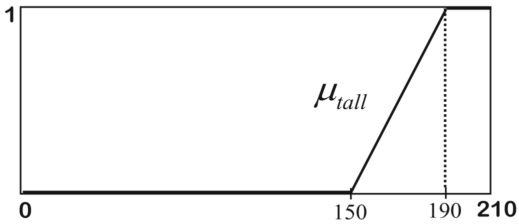
Since μ_{tall} is strictly non-decreasing between 150 and 190, it is $\mu_{tall}(175) < \mu_{tall}(180)$, and $N(\mu_{tall}(180)^2) < N(\mu_{tall}(175)^2)$. Finally,

$$N(N(\mu_{tall}(175)^2)) < N(N(\mu_{tall}(180)^2)), \text{ or } \mu_{tall}(175)^2 < \mu_{tall}(180)^2,$$

and q is strictly more true than p.

Notice that it is not needed to fix S and N, but only a form for μ_{tall}, as well as to accept that $\mu_{very\ P}(x) = \mu_P(x)^2$, $\mu_{mol\ P}(x) = \sqrt{\mu_P(x)}$, and $A(x) = 210 - x$. This last hypotheses is perfectly reasonable since μ_{tall} is non-decreasing, and then the order $\leqslant_{\mu_{tall}}$ is just the order of [0, 210].

Provided we need to know up to which numerical degree p and q do hold, we can use the linear function in the figure,



and fix $N = N_0$. It results

degree up to which p is true $= 1 - (1 - \mu^2_{tall}(175)) = \mu^2_{tall}(175) = 0.391$
degree up to which q is true $= \mu^2_{tall}(180) = 0.563$,

that shows how is q more true than p.

Example 2.2.50 It is known that the algebra with which fuzzy sets must be combined should satisfy the laws

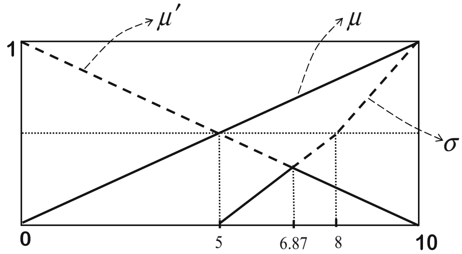
$$\mu + \mu \cdot \sigma = \mu, \quad \mu \cdot (\mu + \sigma) = \mu,$$

as well as that the negation is linear, Determine the triplet (T, S, N) and, with $X =$ [0, 10], and

$$\mu(x) = \frac{x}{10}, \ \sigma(x) = \begin{cases} 0, & \text{if } 0 \leqslant x \leqslant 5 \\ \frac{x-5}{6}, & \text{if } 5 \leqslant x \leqslant 8 \\ \frac{x-6}{4}, & \text{if } 8 \leqslant x \leqslant 10, \end{cases}$$

compute $\mu \cdot \sigma$, $\mu + \sigma$, and $\mu' + \sigma$.

Solution. The first law of absorption $\mu + \mu \cdot \sigma = \mu$, implies $S = \max$ for any T. The second law of absorption $\mu \cdot (\mu + \sigma) = \mu$, implies $T = \min$ for any S. Hence $(T, S) = (\min, \max)$, and the only linear N is $N = N_0$. Hence, $(T, S, N) = (\min, \max, 1 - id)$. With the graphics of μ and σ in the figure,



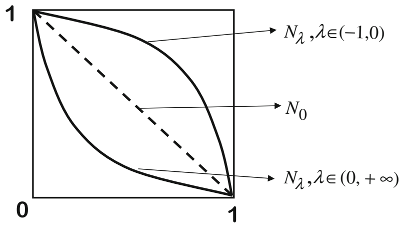
it follows $\mu \cdot \sigma = \min(\mu, \sigma) = \sigma$, $\mu + \sigma = \max(\mu, \sigma) = \mu$, and $\mu' + \sigma = (1-\mu) + \sigma$ is the pointed curve. That is,

$$(\mu' + \sigma)(x) = \begin{cases} \mu'(x), & \text{if } x \in [0, 6.87] \\ \sigma(x), & \text{if } x \in [6.87, 10]. \end{cases}$$

Notice that $x = 6.87$ comes from the equation $1 - \frac{x}{10} = \frac{x-5}{6}$.

Example 2.2.51 Consider the Sugeno's family of strong negations $N_\lambda(x) = \frac{1-x}{1+\lambda x}(\lambda > -1)$. If $-1 < \lambda_1 < \lambda_2$, it follows $1 + \lambda_1 x < 1 + \lambda_2 x$, or $\frac{1}{1+\lambda_2 x} < \frac{1}{1+\lambda_1 x}$, that is, $N_{\lambda_2}(x) < N_{\lambda_1}(x)$. Hence,

If $\lambda \leqslant 0 : N_0 \leqslant N_\lambda$
If $0 \leqslant \lambda : N_\lambda \leqslant N_0$



Compare the graphics of $\mu' = N_0 \circ \mu$ and $\mu' = N_1 \circ \mu$, in a figure, if

$$\mu(x) = \begin{cases} 0, & \text{if } x \in [0, 3] \cup [7, 10] \\ 1, & \text{if } x \in [4, 6] \\ x - 3, & \text{if } x \in [3, 4] \\ 7 - x, & \text{if } x \in [6, 7] \end{cases}$$

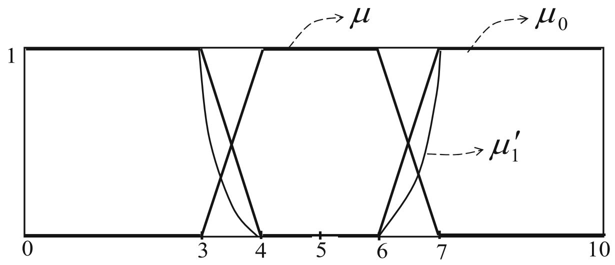
Solution. It is

$$\mu'(x) = \begin{cases} 1, & \text{if } x \in [0, 3] \cup [0, 4] \\ 0, & \text{if } x \in [4, 6] \\ \frac{4-x}{x}, & \text{if } x \in [3, 4] \\ \frac{x-6}{x}, & \text{if } x \in [6, 7], \end{cases}$$

and $\mu'_1(x) = N_1(\mu(x)) = \frac{1-\mu(x)}{1+\mu(x)}$, or

$$\mu'_1(x) = \begin{cases} 0, & \text{if } x \in [0, 3] \cup [7, 10] \\ 1, & \text{if } x \in [4, 5] \\ \frac{4-x}{x-2}, & \text{if } x \in [3, 4] \\ \frac{x-6}{8-x}, & \text{if } x \in [6, 7] \end{cases}$$

Hence,



Look that $N_1(3.2) = 0.\widehat{6}$, $N_1(3.5) = 0.\widehat{3}$, $N_1(3.4) = 0.43$, $N_1(3.6) = 0.25$, $N_1(3.8) = 0.\widehat{1}$, but $N_0(3.2) = 0.8$, $N_0(3.5) = 0.5$, $N_0(3.6) = 0.4$, $N_0(3.4) = 0.6$, , $N_0(3.6) = 0.4$, $N_0(3.8) = 0.2$.

Example 2.2.52 The negation is linear, and the standard algebra must verify the law $\mu = \mu \cdot \sigma + \mu \cdot \sigma'$. Determine the triplet (T, S, N), and with $\mu(x) = \frac{x}{10}$ (in $X = [0, 10]$) and

$$\sigma(x) = \begin{cases} 1, & \text{if } x \in [0, 5] \\ \frac{7-x}{2}, & \text{if } x \in [5, 7] \\ 0, & \text{if } x \in [7, 10] \end{cases}$$

check that $\mu \cdot \sigma + \mu \cdot \sigma' = \mu$, $\mu \cdot \sigma + \mu' \cdot \sigma = \sigma$, and $(\mu \cdot \sigma')' = \sigma + \mu' \cdot \sigma'$.

Solution. From N linear, $N = N_0$, it follows that we can take $\varphi = \mathrm{id}$, and from $\mu \cdot \sigma + \mu \cdot \sigma' = \mu$ it follows $T = \mathrm{prod}_\varphi$, $S = W^*_\varphi$, and $N = N_\varphi$. Hence $(T, S, N) = (\mathrm{prod}, W^*, N)$.

Since $\mu'(x) = 1 - \frac{x}{10}$, and $\sigma'(x) = \begin{cases} 0 \\ \frac{x-5}{2} \\ 1 \end{cases}$, it follows

$$\mu\sigma(x) = \begin{cases} \frac{x}{10} \\ \frac{x(7-x)}{20} \\ 0 \end{cases}, \quad \mu\sigma'(x) = \begin{cases} 0 \\ \frac{x(x-5)}{20} \\ \frac{x}{10} \end{cases}, \quad \mu'\sigma(x) = \begin{cases} 1 - \frac{x}{10} \\ \frac{7-x}{2}(1 - \frac{x}{10}) \\ 0 \end{cases}.$$

Hence,

$$(\mu \cdot \sigma + \mu \cdot \sigma')(x) = \left\{ \begin{array}{l} W^*(\frac{x}{10}, 0) = \frac{x}{10} \\ W^*(\frac{x(7-x)}{20}, \frac{x(x-5)}{20}) = \frac{x}{10} \\ W^*(0, \frac{x}{10}) = \frac{x}{10} \end{array} \right\} = \frac{x}{10} = \mu(x),$$

$$(\sigma \cdot \mu + \sigma \cdot \mu')(x) = \left\{ \begin{array}{l} W^*(\frac{x}{10}, 1 - \frac{x}{10}) = 1 \\ W^*(\frac{x(7-x)}{20}, \frac{7-x}{2}(1 - \frac{x}{10}) = \frac{7-x}{2} \\ W^*(0, 0) = 0 \end{array} \right\} = \sigma(x).$$

Finally, since,

$$(\mu \cdot \sigma')'(x) = 1 - (\mu \cdot \sigma')(x) = \begin{cases} 1 \\ 1 - \frac{x(x-5)}{20} \\ 1 - \frac{x}{10} \end{cases}, \text{ and } (\mu' \cdot \sigma')(x) = \begin{cases} 0 \\ (1 - \frac{x}{10})\frac{x-5}{2} \\ 1 - \frac{x}{10} \end{cases},$$

it results

$$(\sigma + (\mu' \cdot \sigma'))(x) = \begin{cases} W^*(1,0) = 1 \\ W^*(\frac{7-x}{2}, \frac{x-5}{2}(1 - \frac{x}{10})) = 1 - \frac{x(x-5)}{20} = (\mu \cdot \sigma')'(x) \\ W^*(0, 1 - \frac{x}{10}) = 1 - \frac{x}{10} \end{cases}.$$

Example 2.2.53 Predicate F = *high fever* refers to the interval $[37, 42]$ in a clinical thermometer, in which the values $\{37, 37.5, 38, \ldots, 41.5, 42\}$ are significative. Asking an expert one obtains the following fuzzy set

$$\mu_F = 0.3/38.5 + 0.5/39 + 0.7/39.5 + 0.8/40 \\ + 0.9/40.5 + 1/41 + 1/41.5 + 1/42$$

where it is clear that $0/37 + 0/37.5 + 0/38$ is avoided since this values of the body's temperature are not significative for F. With all that, give the membership function of P = *very high fever*, Q = *more or less high fever*, R = *low fever, S* = *not high fever*.

Solution. With the usual definition $\mu_{very\ F} = \mu_F^2$, $\mu_{mol\ F} = +\sqrt{\mu_F}$, $\mu_{low\ F} = \mu_F(37 + 42 - x) = \mu_F(79 - x)$, $\mu_{not\ F} = 1 - \mu_F$, it results:

- $\mu_P = 0.09/38.5 + 0.25/39 + 0.49/39.5 + 0.64/40 + 0.81/40.5 + 1/41 + 1/41.5 + 1/42$.
- $\mu_Q = 0.55/38.5 + 0.7/39 + 0.84/39.5 + 0.89/40 + 0.95/40.5 + 1/41 + 1/41.5 + 1/42$.
- $\mu_R = 1/37 + 1/37.5 + 1/38 + 0.9/38.5 + 0.8/39.5 + 0.7/39.5 + 0.5/40.5 + 0.3/40.5$.
- $\mu_S = 1/37 + 1/37.5 + 1/38 + 0.7/38.5 + 0.5/39.5 + 0.3/39.5 + 0.2/40.5 + 0.1/40.5$.

Notice the incoherence produced by $\mu_S = \mu_{not\ F} \leqslant \mu_{low\ F} = \mu_R$. An incoherence showing that it cannot be taken the representation $\mu_{not\ F} = 1 - \mu_F$, but some $\mu_{not\ F} = N \circ \mu_F$ with $N \geq N_0$.

For example, if $N(x) = \frac{1-x}{1-0.9x}$, it is

$\mu_S = 1/37 + 1/37.5 + 1/38 + +0.96/38.5 + 0.91/39.5 + 0.81/39.5 + 0.71/40 + 0.53/40.5 + 1/41 + 1/41.5 + 1/42$, showing $\mu_{low\ F} \leqslant \mu_{not\ F}$.

Look that

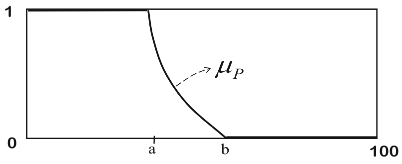
$\mu_{F\ \&\ low\ F} = 0.3/38.5 + 0.5/39.5 + 0.7/39.5 + 0.5/40 + 0.3/40.5$

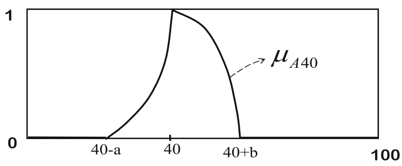
provided $\mu_{F\ \&\ low\ F} = \min(\mu_F, \mu_{low\ F})$.

Example 2.2.54 Describe in fuzzy terms, the statements

$p =$ John is young and around forty,
$q =$ John is old or around forty.

Solution. The solution will come after representing the predicates $P =$ *young*, *aP = Old, A40 = around forty*, in a scale of 0–100 years. The general forms of μ_P, μ_{aP}, and μ_{A40}, are





with $\mu_{aP}(x) = \mu_P(100 - x)$ since μ_P is non-decreasing. Once these functions were established accordingly with the current use of P and $A40$, it will be:

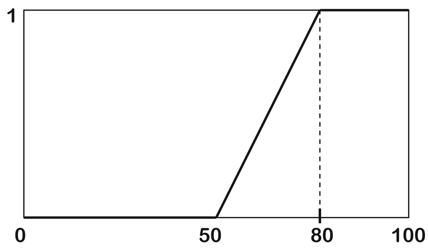
$$\text{Degree}(p) = T(\mu_P(x), \mu_{A40}(x)), \quad \text{Degree}(q) = S(\mu_{aP}(x), \mu_{A40}(x)),$$

with convenient continuous t-norm T and t-conorm S. This formulas are the description of p and q in fuzzy terms.

For example, if $a = 20$, $b = 50$, $\mu_P = \frac{50-x}{30}$, if $20 \leqslant x \leqslant 50$, and $40 - a = 30$, $40 + b = 50$, with μ_{A40} piece-wise linear, $T = \min$, $S = \max$, with

$$\mu_{aP}(x) = \mu_P(100 - x) = \begin{cases} 0, & \text{if } 0 \leqslant x \leqslant 50 \\ \frac{x-50}{30}, & \text{if } 50 \leqslant x \leqslant 80 \\ 1, & \text{if } 80 \leqslant x \leqslant 100, \end{cases}$$

the graphics is



The slashed function describes p, and the continuous one describes q. Of course, $\text{Degree}(p) \leqslant \min(\mu_P(x), \mu_{A40}(x)) \leqslant \mu_P(35) = 0.5$.

Remark 2.2.55 To select T and S, the following points could be taken into account,

- It could be perfectly the case that 'John is young and not young' with a positive degree. Hence, the laws of Non-contradiction can be avoided, and $T \notin \{W\}$.

- Since it is reasonable to accept that 'John is old' *and* 'John is old' does coincide with 'John is old', and 'John is young' *or* 'John is young' does coincide with 'John is young' , we can decide to take either $T = \min$, $S = \max$, or T and S as ordinal-sums.
- Provided idempotency is avoided i.e., 'John is young' *and* 'John is young' does coincide with 'John is very young', instead of $T = \min$, we can take $T = \text{prod}$, that is more interactively than min. In this case, because it does not seem that duality should be avoided, we could take $S = \text{Prod}^*$, and then $\mu_{very\ young\ or\ mol\ old}(45) = \text{Prod}^*(\frac{1}{36}, \frac{\sqrt{6}}{6}) = \frac{1}{6}(1 + \frac{35\sqrt{6}}{36}) = 0.5636$, that is greater than the value $\frac{\sqrt{6}}{6} = 0.408$ obtained with $=$ max.

Example 2.2.56 In $X = [0, 10]$ the predicate $P = \text{big}$ is represented by $\mu(x) = \frac{x}{10}$. In which points in $[0, 10]$ is the degree of 'big' less than that of 'not big'?

Solution. Given μ, the problem is to find for which $x \in X$ it is $\mu(x) \leqslant \mu'(x) = N_\varphi(\mu(x)) = \varphi^{-1}(1 - \varphi(\mu(x)))$, that is, $\mu(x) \leqslant \varphi^{-1}(\frac{1}{2})$. Then,

- If $N = N_0$, $\frac{x}{10} \leqslant 1 - \frac{x}{10}$, or $x \leqslant 5$.
- If $N = N_1$, $\frac{x}{10} \leqslant \frac{1-\frac{x}{10}}{1+\frac{x}{10}}$, or $x^2 + 20x - 100 \leqslant 0$, that means $x \leqslant 10(\sqrt{2} - 1) = 4.142$
- If $N = N_2$, $\frac{x}{10} \leqslant \frac{1-\frac{x}{10}}{1+\frac{2x}{10}}$, or $x^2 + 10x - 50 \leqslant 0$, that means $x \leqslant \sqrt{75} - 5 = 3.66$

Hence,

- If $N = N_0$, the set is $[0, 5]$, and the threshold (of selfcontradiction) of *big* is 5.
- If $N = N_1$, the set is $[0, 4.142]$, and the threshold is 4.142
- If $N = N_2$, the set is $[0, 3.66]$, and the threshold is 3.66

Notice that changing *big* by *not big*, the thresholds do remain but the sets are, respectively, $[5, 10]$, $[4.142, 10]$, and $[3.66, 10]$.

2.3 On Aggregating Imprecise Information

The kind of problems this section will deal with are like the following. An exam is corrected by three referees R_1, R_2, R_3, each one with a different weight of strongness $W(R_i) \in [0, 1]$, $1 \leqslant i \leqslant 3$, such that $\sum_{i=1}^{3} W(R_i) = 1$. Each referee assigns a numerical qualification $p_i \in [0, 10]$ to the exam delivered by a given student. How these qualification can be "aggregated" to obtain final qualification for the student's exam? A recognized usual way of doing it is by the *weighted mean*:

$$\frac{1}{10}Q = \frac{p_1}{10} \cdot W(R_1) + \frac{p_2}{10} \cdot W(R_2) + \frac{p_3}{10} \cdot W(R_3),$$

with $\frac{p_i}{10} \in [0, 10]$. For example, if $W = (0.5, 0.3, 0.2)$ and $P = (7, 6, 5)$, it follows

$$\frac{1}{10}Q = 0.7 \times 0.5 + 0.6 \times 0.3 + 0.2 \times 0.5 = 0.63$$

that implies $Q = 6.3$. Provided the three referees have the same weight, it is $W = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$, and then, $Q = p_1 \cdot \frac{1}{3} + p_2 \cdot \frac{1}{3} + p_3 \cdot \frac{1}{3} = \frac{p_1+p_2+p_3}{3} = \frac{7+6+5}{3} = 6$, is just the arithmetic *mean* of the three qualifications.

Another way of obtaining the final qualification, this time by ignoring the referee's character, is by the *geometric mean*

$$Q = \sqrt[3]{p_1 \cdot p_2 \cdot p_3} = \sqrt[3]{7 \times 6 \times 5} = 5.94,$$

showing that in a problem with $p_1 = p_2 = 10, p_3 = 0$, it results $Q = \sqrt[3]{10 \times 10 \times 0} = 0$, when the arithmetic mean is $\frac{20}{3} = 6.67$.

2.3.1 Aggregation Functions

Most of these problems are "represented" by the so-called *Aggregation Functions*, that is, functions

$$A : [0, 1]^n \to [0, 1],$$

such that

1. A is continuous in all variables
2. $A(0, \ldots, 0) = 0$, and $A(1, \ldots, 1) = 1$
3. If $x_1 \leqslant y_1, \ldots, x_n \leqslant y_n$, then $A(x_1, \ldots, x_n) \leqslant A(y_1, \ldots, y_n)$.

Sometimes it is said that A is an *n-dimensional aggregation function*. Continuous t-norms and continuous t-conorms are 2-dimensional aggregation functions.

Of the many types of aggregation functions, a particular and important type are the *quasi-linear means*,

$$M(x_1, \ldots, x_n) = f^{-1}\left(\sum_{i=1}^{n} p_i . f(x_i)\right)$$

with $(p_1, \ldots, p_n)$ in $[0, 1]$, verifying $\sum\limits_{i=1}^{n} p_i = 1$, and $f : [0, 1] \to \mathbb{R}$, continuous, one-to-one, and monotonic. Function f is called the *generator* of M.

Notice that if f is the identity $f(x) = x$, we get the *weighted means*:

$$M(x_1, \ldots, x_n) = \sum_{i=1}^{n} p_i \cdot x_i,$$

that with $p_1 = \frac{1}{n}$ $(1 \leqslant i \leqslant n)$ is the arithmetic mean

$$M(x_1, \dots, x_n) = \frac{1}{n}\sum_{i=1}^{n} x_i,$$

and with $f(x) = -\log x$, and $p_1 = \frac{1}{n}$ $(1 \leqslant i \leqslant n)$ is the geometric mean

$$M(x_1, \dots, x_n) = \sqrt[n]{p_1 \cdot p_2 \cdots p_n}.$$

With $f(x) = x^\alpha$ $(\alpha > 0)$, is $f^{-1}(x) = x^{\frac{1}{\alpha}}$, and with $p_1 = \frac{1}{n}$, it is obtained the family of quasi-linear means,

$$M_\alpha(x_1, \dots, x_n) = \left(\frac{x_1^\alpha + \cdots + x_n^\alpha}{n}\right)^{\frac{1}{\alpha}}.$$

In particular, with $\alpha = 1$, is M_1 the arithmetic mean, and with $\alpha = -1$, it follows

$$M_{-1}(x_1, \dots, x_n) = \frac{n}{\frac{1}{x_1} + \cdots + \frac{1}{x_n}} \text{ (provided } x_1, \dots, x_n \neq 0),$$

called *Harmonic Mean*. As it is easy to prove,

$$\lim_{\alpha \to 0} M_\alpha(x_1, \dots, x_n) = \sqrt[n]{x_1 \cdots x_n}$$

$$\lim_{\alpha \to \infty} M_\alpha(x_1, \dots, x_n) = \max(x_1, \dots, x_n)$$

$$\lim_{\alpha \to -\infty} M_\alpha(x_1, \dots, x_n) = \min(x_1, \dots, x_n).$$

2.3.2 Ordered Weighted Means

It is said that $M : [0, 1]^n \to [0, 1]$ is a *mean*, when M is continuous, monotonic, and verifies:

$$\min(x_1, \dots, x_n) \leqslant M(x_1, \dots, x_n) \leqslant \max(x_1, \dots, x_n).$$

Since, $\min(0, \dots, 0) \leqslant M(0, \dots, 0) \leqslant \max(0, \dots, 0) = 0$, it results $M(0, \dots, 0) = 0$. Since, $\min(1, \dots, 1) \leqslant M(1, \dots, 1) \leqslant \max(1, \dots, 1) = 1$, it results $M(1, \dots, 1) = 1$. Hence, of course, quasi-linear means are means, but there are more of such means. An important and useful example are the *Ordered Weighted Means* (OWA). Its definition is the following:

$O : [0, 1]^n \to [0, 1]$ is an OWA, if $O(x_1, \ldots, x_n)$ is obtained under the process,

- Select weights $p_1, \ldots, p_n$ in [0, 1], such that $\sum_{i=1}^{n} p_i = 1$.
- Permute the n-pla $(x_1, \ldots, x_n)$, to the n-pla $(x_1^*, \ldots, x_n^*)$ such that $x_1^* \leqslant \cdots \leqslant x_n^*$)
- $O(x_1, \ldots, x_n) = \sum_{i=1}^{n} p_i \cdot x_i^*$.

For example, if $n = 2$,

$$O(x_1, x_2) = p_1 \cdot \min(x_1, x_2) + p_2 \cdot \max(x_1, x_2), \text{ with } p_1 + p_2 = 1.$$

If $n = 4$, and the weights are (0.2, 0.4, 0.3, 0.1), it is

$$\begin{aligned} O(0.2, 0.5, 0.7, 0.3) &= O(0.2, 0.3, 0.5, 0.7) \\ &= 0.2 \times 0.2 + 0.4 \times 0.3 + 0.3 \times 0.5 + 0.1 \times 0.7 = 0.74. \end{aligned}$$

2.3.3 More on Aggregations

Because they are associative, continuous t-norms and continuous t-conorms can be extended to n-dimensional aggregation functions. For example, with $n = 3$,

$$T(x_1, x_2, x_3) = T(x_1, T(x_2, x_3)) = T(T(x_1, x_2), x_3)) = \cdots$$

$$S(x_1, x_2, x_3) = S(x_1, S(x_2, x_3)) = S(S(x_1, x_2), x_3)) = \cdots$$

Nevertheless, not all aggregation functions are associative. For example, if M is the arithmetic mean, $M(x_1, M(x_2, x_3)) = \frac{2x_1+x_2+x_3}{4}$, but $M(M(x_1, x_2), x_3)) = \frac{x_1+x_2+2x_3}{4}$. Concerning means, the only associative are min, and max.

In general, Aggregation Functions are not commutative. For example, a 2-dimensional quasi-linear mean

$$M(x_1, x_2) = f^{-1}(p_1 f(x_1) + p_2 f(x_2)), \quad p_1 + p_2 = 1,$$

is commutative if and only if $p_1 = p_2 = \frac{1}{2}$. Arithmetic and geometric means are commutative, but weighted means in general are not.

If T is a continuous t-norm, and S a continuous t-conorm, the function

$$A(x_1, x_2) = p_1 T(x_1, x_2) + p_2 S(x_1, x_2), \quad p_1 + p_2 = 1$$

is an aggregation function that, since $T \leqslant \min \leqslant \max \leqslant S$, in general is not a mean. The only exception is with $T = \min$, and $S = \max$, as it was said before. For example,

- $A(x_1, x_2) = 0.7x_1.x_2 + 0.3W^*(x_1, x_2)$
- $A(x_1, x_2) = 0.6 \min(x_1, x_2) + 0.4(x_1 + x_2 - x_1.x_2)$
- $A(x_1, x_2) = 0.6W(x_1, x_2) + 0.4 \max(x_1, x_2)$,

are aggregation functions.

2.3.4 Examples

The pointwise aggregation of classical sets is not, in general, a classical set, but a fuzzy one. For example, the arithmetic mean verifies

$$M(0, 0) = 0, \quad M(0, 1) = M(1, 0) = \frac{1}{2}, \quad M(1, 1) = 1$$

and, if A, B are crisp subsets, $M(A, B)$ is not a crisp subset if given by $M(\mu_A, \mu_B)(x) = M(\mu_A(x), \mu_B(x))$. On the contrary, with the geometric mean G, it is

$$G(0, 0) = G(0, 1) = G(1, 0) = 0, \quad G(1, 1) = 1,$$

and $G(A, B)$ is a crisp set.

In all cases, if $\mu \in [0, 1]^X$, $\sigma \in [0, 1]^Y$, and A is an aggregation function, then

$$A(\mu, \sigma)(x, y) = A(\mu(x), \sigma(y)),$$

for all $x \in X$, $y \in Y$, is a fuzzy set $A(\mu, \sigma) \in [0, 1]^{X \times Y}$ called the aggregation of μ and σ. When $X = Y$ it could be defined the fuzzy set $A(\mu, \sigma) \in [0, 1]^X$,

$$A(\mu, \sigma)(x) = A(\mu(x), \sigma(x)), \quad \text{for all } x \in X.$$

Example 2.3.1 If $X = \{1, 2, 3, 4, 5\}$, and $\mu = 0.6/1 + 0.7/2 + 0.5/3 + 1/4$, $\sigma = 0.9/1 + 0.5/3 + 0.7/4 + 0.8/5$, compute $M(\mu, \sigma)$, $G(\mu, \sigma)$, and $O(\mu, \sigma)$ with O the OWA with weights $p_1 = 0.4$, $p_2 = 0.6$.

Solution.

$M(\mu, \sigma) = 0.75/1 + 0.35/2 + 0.5/3 + 0.85/4 + 0.4/5$
$G(\mu, \sigma) = 0.735/1 + 0/2 + 0.5/3 + 0.837/4 + 0/5$
$O(\mu, \sigma) = (0.4 \times 0.6 + 0.6 \times 0.9)/1 + (0.4 \times 0 + 0.6 \times 0.7)/2 + (0.4 \times 0.5 + 0.6 \times 0.5)/3 + (0.4 \times 0.7 + 0.6 \times 1)/4 + (0.4 \times 0 + 0.6 \times 0.8)/5 = 0.72/1 + 0.42/2 + 0.5/3 + 0.88/4 + 0.48/5$.

Notice that $G(\mu, \sigma) \leqslant M(\mu, \sigma)$, but that neither $G(\mu, \sigma)$ and $O(\mu, \sigma)$, nor $M(\mu, \sigma)$ and $O(\mu, \sigma)$, are order-comparable.

Example 2.3.2 A linguistic variable has the fuzzy values $H = high$, $S = short$ and $M = medium$, with H represented by

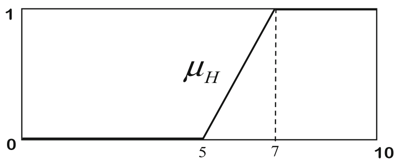
$$\mu_H(x) = \begin{cases} 0, & \text{if } 0 \leqslant x \leqslant 5 \\ 1, & \text{if } 7 \leqslant x \leqslant 10 \\ \frac{x-5}{2}, & \text{if } 5 \leqslant x \leqslant 7. \end{cases}$$

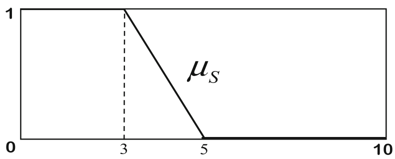
In the two suppositions, $M = H' \cdot S'$, and that M is the aggregation of H and S under the weighted mean $A(x_1, x_2) = 0.3x_1 + 0.7x_2$, compute μ_M.

Solution. Since S is an antonym of H, and μ_H is monotonic, $\mu_S(x) = \mu_H(10 - x)$, that is

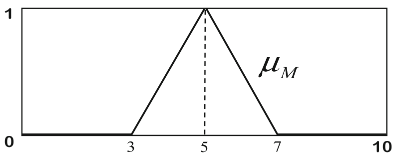
$$\mu_S(x) = \begin{cases} 1, & \text{if } 0 \leqslant x \leqslant 3 \\ 0, & \text{if } 5 \leqslant x \leqslant 10 \\ \frac{5-x}{2}, & \text{if } 3 \leqslant x \leqslant 5. \end{cases}$$

The solution for the first supposition appears in the following sequence of figures (with $\cdot = \min, ' = 1 - \mathrm{id}$).





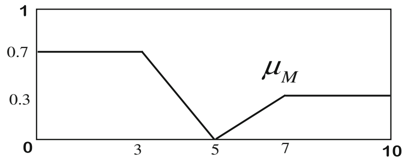
$$\Downarrow$$



The solution for the second supposition is

$$\mu_H(x) = \mu_{A(H,S)}(x) = \begin{cases} 0.7, & \text{if } 0 \leqslant x \leqslant 3 \\ \frac{0.7(5-x)}{2}, & \text{if } 3 \leqslant x \leqslant 5 \\ \frac{0.3(x-5)}{2}, & \text{if } 5 \leqslant x \leqslant 7 \\ 0.3, & \text{if } 7 \leqslant x \leqslant 10, \end{cases}$$

graphically,



Chapter 3
Reasoning and Fuzzy Logic

3.1 What Does It Mean "Logic"?

The question is, in fact, a philosophical one whose discussion does not correspond to this text, and that received a lot of comments and discussions by philosophers. Instead of such question, there is the more particular,what is *a logic*?, that can be answered not philosophicaly but in terms of the mathematical definition of what is a consequence's operator. A definition that corresponds to an abstraction of the term "deduction".

3.1.1 Logic and Consequence Operators

A logic is a triplet $(X, \mathcal{A}, C)$, where $\mathcal{A} \subset \mathbb{P}(X)$ is a family of parts of X, and $C : \mathcal{A} \to \mathcal{A}$ is a mapping verifying

1. For all $P \in \mathcal{A}$, $P \subset C(P)$.
2. If $P, Q \in \mathcal{A}$, and $P \subset Q$, then $C(P) \subset C(Q)$.
3. For all $P \in \mathcal{A}$, $C(C(P)) \subset C(P)$.

It can be said that the pair $(\mathcal{A}, C)$ defines a logic in the set X.

From 1. and 2., it follows $C(P) \subset C(C(P))$, and from 3. results

3′. For all $P \in \mathcal{A}$, $C(C(P)) = C(P)$, or $C^2 = C$.

Hence, a consequence's operator is one that is extensive (1.), monotonic (2.), and a closure (3′.). These operators were introduced by the logician Alfred Tarski in the thirties of 20th century, and are called to be *compact* provided for all $P \in \mathcal{A}$, it exists a finite set $\{p_1, \ldots, p_n\} \subset P$ such that

4. $C(P) = C\{p_1, \ldots, p_n\}$.

E. Trillas and L. Eciolaza, *Fuzzy Logic*, Studies in Fuzziness and Soft Computing 320, DOI 10.1007/978-3-319-14203-6_3

It is clear that if P is finite, C is compact. This happens, for example if X is finite.

A consequence's operator C is *consistent* when '$q \in C(P) \Rightarrow q' \notin C(P)$'.

Example 3.1.1 In a finite lattice $(X, \cdot, +; 0, 1)$, let's consider $\mathcal{A} = \mathbb{P}_0(X) = \{P = \{p_1, \ldots, p_n\} \subset X, p_\wedge = p_1 \cdots p_n \neq 0\}$, and the operator

$$Cons : \mathbb{P}_0(X) \to \mathbb{P}_0(X), \text{ defined by } Cons(P) = \{q \in X; p_\wedge \leqslant q\},$$

with the partial order $\leqslant$ given by $x \leqslant y \Leftrightarrow x \cdot y = x$. *Cons* is a consequence's operator:

1. Since $p_\wedge \leqslant p_i$, it is $p_i \in Cons(p)$, for all $p_i \in P$. That is, $P \subset Cons(P)$.
2. If $P = \{p_1, \ldots, p_n\} \subset \{p_1, \ldots, p_n, p_{n+1}, \ldots, p_m\} = Q$, since $q_\wedge = p_1 \cdots p_n \cdot p_{n+1} \cdots p_m \leqslant p_1 \cdots p_n = p_\wedge$, if $q \in Cons(P)$, from $p_\wedge \leqslant q$, it follows $q_\wedge \leqslant q$, and $q \in Cons(Q)$. Hence, $P \subset Q$ implies $Cons(P) \subset Cons(Q)$.
3. Obviously, Min $Cons(P) = p_\wedge$, hence, $Cons(P) \in \mathbb{P}_0(X)$ since Min $Cons(P) \neq 0$. Then, if $q \in Cons(Cons(P))$, or Min $Cons(P) \leqslant q$, or $p_\wedge \leqslant q$, it follows $q \in Cons(P)$. Hence, $Cons(Cons(P)) \subset Cons(P)$.

In any finite lattice $(X, \cdot, +; 0, 1)$, it can be considered the logic $(X, \mathbb{P}_0(X), Cons)$, and it can be proved that if the lattice is endowed with a complement $'$ such that $(X, \cdot, +, '; 0, 1)$ is a Boolean algebra, any operator of consequences $C : \mathbb{P}_0(X) \to \mathbb{P}_0(X)$ verifies $C \subset Cons$, that is $C(P) \subset Cons(P)$, for all $P \in \mathbb{P}_0(X)$. *Cons* is the biggest operator of consequences in a Boolean algebra with $\mathcal{A} = \mathbb{P}_0(X)$.

The set of *premises* P is *consistent*, if $p_\wedge$ is not self contradictory ($p_\wedge \nleqslant p'_\wedge$). If it were $p_\wedge \leqslant p'_\wedge$, it will be also $p'_\wedge \in Cons(P)$, that is absurd if *Cons* is consistent. In this cases, $q \in Cons(P)$ and $q' \in Cons(P)$, or $p_\wedge \leqslant q$ and $p_\wedge \leqslant q'$ (or $q \leqslant p'_\wedge$) implies $p_\wedge \leqslant p'_\wedge$, that is absurd. Thus, $q \in Cons(P)$ does imply $q' \notin Cons(P)$, and the operator *Cons* is consistent.

Remark 3.1.2 Instead of a lattice, let us take the set $[0, 1]^X$ endowed with a fuzzy intersection

$$\mu_\wedge = \mu_1 \cdots \mu_n = T \circ (\mu_1 \times \cdots \times \mu_n)$$

(T a continuous t-norm), the partial order $\mu \leqslant \sigma \Leftrightarrow \mu(x) \leqslant \sigma(x)$, for all $x \in X$, and the empty set $\mu_0 = \mu_\emptyset$. Take the set $\mathbb{P}_0([0, 1]^X)$ that consists of the finite subsets $P = \{\mu_1, \ldots, \mu_n\} \subset [0, 1]^X$ such that $\mu_\wedge \neq \mu_0$. The definition

$$Cons(P) = \{\sigma \in [0, 1]^X; \mu_\wedge \leqslant \sigma\},$$

allows the same result as in the case before.

There is a, perhaps alternative, way of constructing a logic in a set X. It follows from the following results.

- If $(X, \mathcal{A}, C)$ is a logic in X, the binary relation '$x \leqslant_C y \Leftrightarrow y \in C(\{x\})$', defined if $\{x\} \in \mathcal{A}$, is a preorder.

Proof Since $\{x\} \subset C(\{x\})$, it is $x \leqslant_C x$, for all $\{x\} \in \mathcal{A}$. If $x \leqslant_C y$ and $y \leqslant_C z$, it is $y \in C(\{x\})$, and $z \in C(\{y\})$, hence from $\{y\} \subset C(\{x\})$ follows $C(\{y\}) \subset C(C(\{x\})) = \{x\}$, and $z \in C(\{x\})$, or $x \leqslant_C z$. □

Notice that the preorder $\leqslant_C$ is defined only with the pairs $(x, y) \in X \times X$ such that $\{x\} \in \mathcal{A}$ and $\{y\} \in \mathcal{A}$.

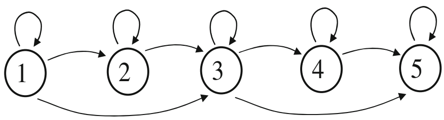
- Given a set X and $\mathcal{A} \subset \mathbb{P}(X)$, if $\leqslant$ is a preorder in X such that

$$\text{If } x \in P, \text{ and } x \leqslant y, \text{ it is } y \in Q, \text{ for some } Q \in \mathcal{A},$$

the operator $C_\leqslant : \mathcal{A} \to \mathcal{A}$, defined by $C_\leqslant(P) = \{y \in X; \exists x \in P \ \& \ x \leqslant y\}$, verifies,

1. $\forall x \in P$, it is $x \leqslant x$, then $x \in C_\leqslant(P) : P \subset C_\leqslant(P)$
2. If $P \subset Q$, and $y \in C_\leqslant(P)$, there is $x \in P$ such that $x \leqslant y$. Since $x \in Q$, it is $y \in C_\leqslant(Q)$. Hence, $C_\leqslant(P) \subset C_\leqslant(Q)$.
3. If $y \in C_\leqslant(C_\leqslant(P))$, there is $x \in C_\leqslant(P)$, such that $x \leqslant y$. But there is also $z \in P$ such that $z \leqslant x$. Hence $z \leqslant y$, and $y \in C_\leqslant(P)$. That is, $C_\leqslant(C_\leqslant(P)) \subset C_\leqslant(P)$.
4. The preorder $\leqslant_{C_\leqslant}$ coincides with the initial $\leqslant$, since $x \leqslant_{C_\leqslant} y \Leftrightarrow y \in C_\leqslant(\{x\}) \Leftrightarrow x \leqslant y$.

Example 3.1.3 Let it be the set $X = \{1, 2, 3, 4, 5\}$ endowed with the preorder



It is, for example $C_\leqslant(\{1, 5\}) = \{1, 5, 2, 3, 4\}$, $C_\leqslant(\{4, 5\}) = \{4, 5\}$, $C_\leqslant(\{2, 3, 4\}) = \{2, 3, 4, 5\}$, $C_\leqslant(\{1, 2, 4, 5\}) = \{1, 2, 4, 5, 3\} = X$, $C_\leqslant(\{5\}) = \{5\}$, and $C_\leqslant(\{3\}) = \{3, 4, 5\}$.

Hence, it is possible to identify a logic in a set with a preordering of it.

3.1.2 Conjecturing

A process to pass from a set of premises $P = \{\mu_1, \ldots, \mu_n\}$ to a 'conclusion' σ is a *conclusive reasoning*, that is sometimes symbolized by $P \vdash \sigma$. A conclusive reasoning $P \vdash \sigma$ is *deductive* if there exists either an operator of consequences C, or a preorder $\leqslant$, such that $P \vdash \sigma$ is equivalent to $\sigma \in C(P)$, or to $\mu_i \leqslant \sigma$ for all $\mu_i \in P$.

Anyway, not all conclusive reasonings are deductive. In common reasoning if there is the possibility of stating $P \vdash \sigma$ and defining $C(P) = \{\sigma; P \vdash \sigma\}$, the axiom of monotonicity is not always verified by such C, but it could verify

- If $P \subset Q$, then $C(Q) \subset C(P)$, C is anti-monotonic.
- If $P \subset Q$, then $C(P)$ and $C(Q)$ are not comparable, that is, it is neither $C(P) \subset C(Q)$, nor $C(Q) \subset C(P)$, C is non-monotonic.

Most common conclusive reasonings are not deductive, but of a conjectural type, that is, the conclusion σ is provisionally accepted because, simply, it is non contradictory with all or with part of the information. It should be noticed that the lost of the monotonicity of C motivates the lost of the transitive property of the preorder $\leqslant_C$.

We will say that $P \vdash \sigma$ is a conjectural kind of conclusive reasoning if there exists an operator of consequences C, such that

$$P \vdash \sigma \Leftrightarrow \sigma' \notin C(P) \Leftrightarrow N \circ \sigma \notin C(P),$$

for some strong negation N. Analogously, if what we have is a preorder $\leqslant$ (instead of C), σ is a conjecture of the information contained in P, when

$$P \vdash \sigma \Leftrightarrow \forall \mu \in P : \mu \nleqslant \sigma'.$$

In what follows, we will only take into account the operator $Cons(P) = \{\sigma; \mu_\wedge \leqslant \sigma\}$, provided $\mu_\wedge \neq \mu_0$. Consequently, the set of the conjectures that, through $Cons$, is associated to any finite set $P = \{\mu_1, \ldots, \mu_n\}$ of premises such that $\mu_\wedge = \mu_1 \cdots \mu_n = T \circ (\mu_1 \times \ldots \times \mu_n) \neq \mu_0$, for some continuous t-norm T, is

$$Conj(P) = \{\sigma \in [0, 1]^X; \mu_\wedge \leqslant \sigma'\}^c,$$

with $\sigma' = N \circ \sigma$, for some strong negation N.

Notice that it is necessary to take a pair of connectives (T, N) to define $Conj(P)$.

Remark 3.1.4 Instead of $\mu_\wedge \neq \mu_0$, in what follows we will suppose that $\mu_\wedge$ is not self-contradictory, that is $\mu_\wedge \nleqslant \mu'_\wedge$. With this hypothesis, if $\sigma \in Cons(P) \Leftrightarrow \mu_\wedge \leqslant \sigma$, it can't be $\mu_\wedge \leqslant \sigma'$ because of $\sigma' \leqslant \mu'_\wedge$ implies $\mu_\wedge \leqslant \mu'_\wedge$. Hence, it should be $\mu_\wedge \nleqslant \sigma'$ or $\sigma \in Conj(P)$. That is, $Cons(P) \subset Conj(P)$. In addition; $Cons$ is consistent. Notice that $\mu_0 \notin Conj(P)$.

In fact, $\mu_\wedge \nleqslant \mu'_\wedge$ is more general than $\mu_\wedge \neq \mu_0$, since $\mu_\wedge = \mu_0$ implies $\mu_\wedge = \mu_0 \leqslant \mu_1 = \mu'_0 = \mu'_\wedge$. Hence, if it is $\mu_\wedge \nleqslant \mu'_\wedge$, it is also $\mu_\wedge \neq \mu_0$.

With this change,

$$Conj(P) = Cons(P) \cup Hyp(P) \cup Sp(P),$$

where

- $Hyp(P) = \{\sigma \in Conj(P); \mu_0 < \sigma < \mu_\wedge\}$, is the set of hypotheses of P.
- $Sp(P) = \{\sigma \in Conj(P); \mu_\wedge \; NC \; \sigma\}$, is the set of speculations of P, where NC means non-comparable under $\leqslant$,

and verifying

$$Cons(P) \cap Hyp(P) = Cons(P) \cap Sp(P) = Hyp(P) \cap Sp(P) = \emptyset.$$

Notice that

- *Consequences follow* (in the partial order $\leqslant$) from $\mu_\wedge$: *all the premises explain the consequences.*
- *Hypotheses explain all the premises*, since $\sigma < \mu_\wedge \leqslant \mu_i$, all the premises follow from each hypothesis.
- Speculations, are the conjectures for which it is neither $\mu_\wedge \leqslant \sigma$, nor $\sigma < \mu_\wedge$.

With all that, processes,

- $P \vdash \sigma$, with $\sigma \in Conj(P)$, is a guessing, or conjectural reasoning
- $P \vdash \sigma$, with $\sigma \in Cons(P)$, is a deduction, or deductive reasoning
- $P \vdash \sigma$, with $\sigma \in Hyp(P)$, is an abduction, or abductive reasoning
- $P \vdash \sigma$, with $\sigma \in Sp(P)$, is an speculation, or speculative reasoning

Concerning this types of conclusive reasonings, it should be pointed out what follows.

1. The set $Conj(P)$ is not always in $A = \mathbb{P}_0(X)$. Hence, it can't be taken as a set of premises, it has not sense to consider $Conj(Conj(P))$, or $Cons(Conj(P))$.
2. If $P \subset Q$, it is $Conj(Q) \subset Conj(P)$, because if $\sigma \in Conj(Q)$, from $Inf\, Q \leqslant Inf\, P$, if $Inf\, Q \nleqslant \sigma'$, it is $Inf\, P \nleqslant \sigma'$. Hence, if $\sigma \in Conj(Q)$ it is $\sigma \in Conj(P)$, and $Conj$ is anti-monotonic.
3. The sets $Hyp(P)$ and $Sp(P)$ are not always in $\mathbb{P}_0(X)$. Hence, they can't be taken as sets of premises.
4. If $P \subset Q$, it is $Hyp(Q) \subset Hyp(P)$, since $\sigma \in Hyp(Q)$, or $\mu_0 < \sigma < Inf\, Q$ implies $\mu_0 < \sigma < Inf\, P$. Hence, the operator Hyp is anti-monotonic.
5. Concerning the operator Sp, if $P \subset Q$ it can be $Sp(P) \not\subseteq Sp(Q)$, and $Sp(Q) \not\subseteq Sp(P)$. Hence, Sp is a non-monotonic operator.

3.2 Reasoning with Conditionals: Representation

3.2.1 What is a Conditional?

A conditional is an statement of the form 'If a, then b' $:= a \to b$, with two previous statements a, b. For example, "If it is raining, then I take an umbrella", where $a =$ It is raining, $b =$ I take an umbrella, or "If the food is well cooked and well served, and the wine is of good quality, then the tip will be higher than usual", wit $a =$ The food is well cooked and well served, and the wine is of good quality, $b =$ The tip will be higher than usual.

Notice that the first example is a crisp conditional, but the second is an imprecise one. In what follows we will take into account the *representation of imprecise conditionals* of the type

$$\text{If } x \text{ is } P, \text{ then } y \text{ is } Q,$$

where $x \in X$, $y \in Y$, P is a predicate (precise or imprecise) in X, and Q is a predicate (precise or imprecise) in Y. For example, with $X = [0, 1]$, $Y = [0, 10]$,

$$\text{If } x \text{ is small, then } y \text{ is big,}$$

or,

$$\text{If } x \text{ is small and } y \text{ is big, then } z \text{ is not small,}$$

with $x, y \in [0, 1]$ and $z \in [0, 10]$.

What it means to represent a conditional statement like "If x is P, then y is Q"? It means to translate it in fuzzy terms. For example, 'x is P' and 'y is Q' will be translated by $\mu_P(x)$ and $\mu_Q(y)$ with adequate fuzzy sets $\mu_P, \mu_Q \in [0, 1]^X$, adequate in the sense that they capture the use of P on X and Q on Y.

But how to represent the full statement "If x is P, then y is Q":= $\mu_P(x) \to \mu_Q(y)$? It is always done, in fuzzy logic, by means of a function $J : [0, 1] \times [0, 1] \to [0, 1]$, such that

$$\mu_P(x) \to \mu_Q(y) = J(\mu_P(x), \mu_Q(y)) \in [0, 1]$$

for all $x \in X$, and $y \in Y$. But, which function J can be taken? It depends on the 'meaning' of the conditional statement, and this requires to look at what happens in general.

Remark 3.2.1 Imprecise conditionals are very useful in ordinary life, for instance,

- If the turn is not so far, and the car's speed is not high, then press softly the break,
- If the the food was well-cooked and of quality, and the service was good, the the tip should be higher than 15 %.

Both in common life and in technology, a lot of imprecise conditionals (rules) are considered. The need for its representation will be obvious in the next section.

3.2.2 The Case of Boolean Algebras

Let us consider, in the first place, the case in which the statements are crisp and belong to a Boolean algebra $(B, \cdot, +, '; 0, 1)$, where $\cdot, +, '$ stand, respectively, for the intersection (and), the union (or), and the complement (negation, not), 0 is the minimum element, and 1 is the maximum. As it is well known, this Boolean algebra is naturally ordered by means of the partial order

$$a \leqslant b \Leftrightarrow a \cdot b = a \Leftrightarrow a + b = b \Leftrightarrow b = a + a' \cdot b.$$

Representing 'If a, then b' by $a \to b$, what is important is that from the set of premises $P = \{a, a \to b\}$ should follow the statement b *as a logical consequence*.

For this goal, it should be $a \cdot (a \to b) \neq 0$, and $a \cdot (a \to b) \leqslant b$. In a boolean algebra it is

$$a \cdot z \leqslant b \Leftrightarrow z \leqslant a' + b,$$

because of:

1. $a \cdot z \leqslant b \Rightarrow a' + a \cdot z = a' + z \leqslant a' + b$, and since $z \leqslant a' + z$, follows $z \leqslant a' + b$
2. $z \leqslant a' + b \Rightarrow a \cdot z \leqslant a \cdot (a' + b) = a \cdot b \leqslant b$.

Then, $a \cdot (a \to b) \leqslant b \Leftrightarrow a \to b \leqslant a' + b$.

A conditional function is a mapping $\to: B \times B \to B$, such that $a \cdot (a \to b) \leqslant b$ for all $a, b \in B$. Hence, in a Boolean algebra, the biggest conditional is $a' + b$, the so-called material conditional, and any smaller function is also a conditional. For example, from

$$a \cdot b \leqslant b \leqslant a' + b, \ a' \leqslant a' + b$$

it follows that $a \to b = a \cdot b, a \to b = b, a \to b = a'$, are conditionals. Analogously, from

$$a' \cdot b' + a \cdot b \leqslant a' + b,$$

it also follows that $a \to b = a' \cdot b' + a \cdot b$ is a conditional. Different ways of writing $a' + b$ in a boolean algebra, are

$$a' \cdot (b + b') + a \cdot b, \ a' + a \cdot b, \ b + a' \cdot b'$$

since $a' \cdot (b + b') + a \cdot b = a' + a \cdot b = (a + a') \cdot (a' + b) = a' + b$, and $b + a' \cdot b' = (b + a') \cdot (b + b') = a' + b$.

Notice that from $z_1 \leqslant a' + b, z_2 \leqslant a' + b$, follows $z_1 \cdot z_2 \leqslant (a' + b) \cdot (a' + b) = a' + b$, $z_1 + z_2 \leqslant (a' + b) + (a' + b) = a' + b$, hence, the union and the intersection of conditionals is also a conditional. For example, $a \cdot b + a' + b = a' + b$ is obviously a conditional.

There are two-variable functions $a \to b$ such that $a \to b \leqslant a' + b$, but are not expressible as a single formula with the connectives $'$, $\cdot$, $+$ as the before considered cases. For example,

$$a \to b = \begin{cases} a \cdot b, & \text{if } a \cdot b \neq 0 \\ a' + b, & \text{if } a \cdot b = 0, \end{cases}$$

verifies $(a \to b) \cdot a = \left\{ \begin{array}{ll} a \cdot b, & \text{if } a \cdot b \neq 0 \\ a \cdot (a' + b) = a \cdot b, & \text{if } a \cdot b = 0, \end{array} \right\} = a \cdot b \leqslant b$, that is, is a conditional. Analogously,

$$a \to b = \begin{cases} 1, & \text{if } a \leqslant b \\ b, & \text{otherwise}, \end{cases}$$

verifies $(a \to b) \cdot a = \left\{ \begin{array}{ll} a, & \text{if } a \leqslant b \\ a \cdot b, & \text{otherwise} , \end{array} \right\} \leqslant b$, that is, is a conditional.

Remark 3.2.2 If the Boolean algebra B is complete, that is, for any $A \subset B$, $A \neq \emptyset$, it exists $SupA \in B$, then

$$Sup\{z \in B; a \cdot z \leqslant b\} = Sup\{z \in B; z \leqslant a' + b\} = a' + b.$$

Remark 3.2.3 The character of conditional of $a'+b$ is exclusive of Boolean algebras. That is, in any ortholattice, the validity of $a \cdot (a' + b) \leqslant b$, for all a, b, forces the ortholattice to be a Boolean algebra.

Remark 3.2.4 $a \to b \leqslant a'+b$, is a property that only holds in Boolean algebras, that is, in ortholattices the equivalence $a \cdot z \leqslant b \Leftrightarrow z \leqslant a' + b$, is not valid. It only holds in Boolean algebras. For example, in orthomodular lattices, both $a \to_1 b = a'+a \cdot b$, and $a \to_2 b = b' + a' \cdot b'$ (that verify $a \to_2 b = b' \to_1 a'$), are conditionals, but is neither $a \to_1 b \leqslant a \to_2 b$ nor $a \to_2 b \leqslant a \to_1 b$.

The conditional $a \to_1 b = a' + a \cdot b$ is called the *Sasaki hook*, and $a \to_2 b = b' + a' \cdot b'$ is the *Dishkant hook*, and, of course, only in Boolean algebras are both coincidental with $a' + b$. The Sasaki and the Dishkant hooks are used as models for the conditional statements in the reasoning in Quantum Logic.

Remark 3.2.5 The scheme of *Modus Ponens*

$$\begin{array}{l} \text{If } a, \text{ then } b \\ a \\ \hline \qquad\qquad b, \end{array}$$

corresponds to *forwards reasoning*, that is, goes from the antecedent a to the consequent b thanks to the conditional $a \to b$, through $a \cdot (a \to b) \leqslant b$. *Backwards reasoning* goes from the consequent to the antecedent (also thanks to $a \to b$), ad it is modeled by the *Modus Tollens* scheme.

$$\begin{array}{l} \text{If } a, \quad \text{then } b \\ \qquad\qquad \text{not } b \\ \hline \text{not } a, \end{array}$$

that is translated by $b' \cdot (a \to b) \leqslant a' \Leftrightarrow a \to b \leqslant (b')' + a' = a' + b$. Thus, in Boolean algebras, $a \to b = a' + b$, also allows backwards reasoning, provided $b' \cdot (a \to b) = b' \cdot (a' + b) = a' \cdot b' \neq 0$, or $a + b \neq 1$. Nevertheless, although $b' \cdot (a \cdot b) = 0 \leqslant a'$, it is clear that the conjunctive conditional $a \to b = a \cdot b$ does not allow backwards reasoning since $b' \cdot (a \to b) = 0$.

3.2.3 Fuzzy Conditionals

Let us return to the case of fuzzy logic, that is, to a conditional linguistic expression, or rule, like 'If x is P, then y is Q', represented in fuzzy terms by

$$(\mu_P \to \mu_Q)(x, y) = J(\mu_P(x), \mu_Q(y)),$$

for all $x \in X$, $y \in Y$. The problem is which function $J : [0, 1] \times [0, 1] \to [0, 1]$, is to be taken at each case if, of course, it gives a conditional. That is, if from the premises {'x is P', 'If x is P, then y is Q'} follows 'y is Q' as a logical consequence. Formally speaking, it should exist a continous t-norm T_0 such that

$$T_0(\mu_P(x), J(\mu_P(x), \mu_Q(y))) \leqslant \mu_Q(y)$$

for all $x \in X$, $y \in Y$. This condition, that should hold for any $\mu_P(x) \in [0, 1]$, and any $\mu_Q(y) \in [0, 1]$, conducts to the inequality

$$T_0(a, J(a, b)) \leqslant b$$

for all a, b in [0, 1], J should verify to represent conditional statements. It is called the *Modus Ponens Inequality*, since it allows the scheme of reasoning

$$\begin{array}{l} \text{If x is P, then y is Q} \\ \text{x is P} \\ \hline \hfill \text{y is Q} \end{array}$$

called the *scheme of Modus Ponens*. J is called a T_0-conditional function (shortly, T_0-conditional).

There is a theorem, whose proof will be omitted, showing that being T_0 a continuous t-norm, it is

$$T_0(a, J(a, b)) \leqslant b \Leftrightarrow J(a, b) \leqslant J_{T_0}(a, b) = \sup\{z \in [0, 1]; T_0(z, a) \leqslant b\}.$$

Hence, for each T_0, the greatest T_0-conditional is the function J_{T_0}, since, it verifies the MP-inequality $T_0(a, J_{T_0}(a, b)) = min(a, b) \leqslant b$.

Remark 3.2.6 For reasons that will be latter on presented, T-conditionals J_T are called R-implications (R shorting residuated). They come directly from the Boolean equation $a' + b = \sup\{z; a \cdot z \leqslant b\}$.

If $\mu, \sigma \in \{0, 1\}^X$, take $(\mu \to \sigma)(x, y) = J_T(\mu(x), \sigma(y)) = \sup\{z \in [0, 1]; T(z, \mu(x)) \leqslant \sigma(y)\}$. If $\mu(x) \in \{0, 1\}$, $\sigma(y) \in \{0, 1\}$, it is

$$J_T(\mu(x), \sigma(y)) = \begin{cases} J_T(0, 0) = 1 \\ J_T(0, 1) = 1 \\ J_T(1, 0) = 0 \\ J_T(1, 1) = 1, \end{cases}$$

that coincides with the values of $\max(1 - \mu(x), \sigma(y))$. That is, all R-implications do coincide with the Boolean material conditional $\mu' + \sigma$ in the case that μ and σ

are crisp sets. R-implications generalize the material conditional. Of course, this will happen with any J such that

$$J(0,0) = 1, J(0,1) = 1, J(1,0) = 0, J(1,1) = 1$$

Example 3.2.7 The immediate generalization of the Boolean conditional $a \to b = a' + b$, is given by $(\mu' + \sigma)(x, y) = \mu'(x) + \sigma(y)$, that is, by $J(a,b) = S(N(a), b)$, for all a, b in [0, 1]. These operators are called S-implications (S shortens 'strong'). With,

- $S = \max$, $N = N_0$, is $J(a,b) = \max(1-a, b)$, called the Kleene-Diennes conditional.
- $S = prod^*$, $N = N_0$, is $J(a,b) = 1-a+ab$, called the Reichenbach conditional.
- $S = W^*$, $N = N_0$, is $J(a,b) = \min(1, 1-a+b)$, called the Łukasiewicz conditional.

Notice that the MP inequality $T_0(S(N(a), b)) \leqslant a$, is verified on the last three cases with $T_0 = W$:

- $W(a, \max(1-a, b)) = \max(0, a+b-1) = W(a,b) \leqslant b$
- $W(a, 1-a+a \cdot b) = \max(0, a \cdot b) = a \cdot b \leqslant b$
- $W(a, \min(1, 1-a+b)) = \max(0, \min(a,b)) = \min(a,b) \leqslant b$,

hence, the three cases are W-conditionals.

Example 3.2.8 Let us see how is J_T, when T is, respectively, the continuous t-norm min, $prod_\varphi$, W_φ.

- $T = \min$, $J_{\min}(a,b) = \sup\{z \in [0,1]; \min(z,a) \leqslant b\} = \begin{cases} 1, & \text{if } a \leqslant b \\ b, & \text{if } a > b \end{cases}$ (*Gödel* implication).
- $T = prod_\varphi$, $J_T(a,b) = \sup\{z \in [0,1]; \varphi(a).\varphi(z) \leqslant \varphi(b)\} = \begin{cases} 1, & \text{if } a \leqslant b \\ \varphi^{-1}(\frac{\varphi(b)}{\varphi(a)}), & \text{if } a > b \end{cases}$ (Goguen implication).
- $T = W_\varphi$, $J_T(a,b) = \sup\{z \in [0,1]; \varphi^{-1}(W(\varphi(a), \varphi(z)) \leqslant b\} = \varphi^{-1}(\min(1, 1 - \varphi(a) + \varphi(b)))$ (Łukasiewicz implication).

Since each J_T is a T-conditional, *Gödel*'s is a min-conditional, Goguen's are prod$_\varphi$-conditionals, and Łukasiewicz's are W$_\varphi$-conditionals. Notice that the S-implications of the form

$$W^*_\varphi(N_\varphi(a), b) = \varphi^{-1}(W^*(\varphi(N_\varphi(a)), \varphi(b))) = \varphi^{-1}(W^*(1-\varphi(a), \varphi(b))) =$$

$$\varphi^{-1}(\min(1, 1-\varphi(a)+\varphi(b))),$$

are exactly the Łukasiewicz's R-implications: the only R-implications that are S-implications are the Łukasiewicz's ones. If, for instance, it were

$$J_{\min}(a,b) = S(N(a),b) = \begin{cases} 1, & \text{if } a \leqslant b \\ b, & \text{if } a > b \end{cases}$$

it will result

$$S(a,b) = \begin{cases} 1, & \text{if } N(a) \leqslant b \\ b, & \text{if } N(a) > b \end{cases}$$

a function that is not a t-conorm, since $S(a,0) = 0 \neq a$, if $a > 0$. Hence, $J_{\min}$ is not an S-implication. An analogous reasoning shows that J_{prod_φ} are not S-implications.

Example 3.2.9 The protoform $\mu \to \sigma = \mu' + \mu \cdot \sigma$ (coming from the Sasaki hook), gives

$$J_1(a,b) = S(N(a), T(a,b)),$$

with S a continuous t-conorm, T a continous t-norm, an N an strong negation. These function are called Q-conditionals (Q for Quantum). For example,

- $S = \max$, $T = \min$, $N = N_0$, is $J_1(a,b) = \max(1-a, \min(a,b))$, is the so-called Early-Zadeh operator
- $S = \max$, $T = prod$, $N = N_0$, is $J_1(a,b) = \max(1-a, ab)$
- $S = prod^*$, $T = \min$, $N = N_0$, is $J_1(a,b) = 1 - a + a^2 b$
- $S = W^*$, $T = W$, $N = N_0$, is $J_1(a,b) = \max(1-a, b)$, that coincides with the Kleene-Diennes implication
- $S = W^*$, $T = prod$, $N = N_0$, is $J_1(a,b) = 1 - a + ab$, that coincides with the Reichenbach implication
- $S = W^*$, $T = \min$, $N = N_0$, is $J_1(a,b) = \min(1, 1-a+b)$, that coincides with the Łukasiewicz implication

With which t-norm T_0 do verify the MP inequality these Q-operators? For instance,

- $W(a, \max(1-a, \min(a,b))) = \max(0, a + \min(a,b) - 1) = W(a, \min(a,b)) \leqslant \min(a,b) \leqslant b$
- $W(a, \max(1-a, a\cdot b)) = \max(0, a + a\cdot b) - 1) = W(a, a\cdot b) \leqslant a\cdot b \leqslant b$
- $W(a, 1 - a + a^2 b) = \max(0, a^2, b) = a^2 b \leqslant b$
- $W(a, \max(1-a, b)) \leqslant b$ (as it is proven before)
- $W(a, 1 - a + ab) \leqslant b$ (as it is proven before)

Example 3.2.10 The protoform $\mu \to \sigma = \sigma + \mu' \cdot \sigma'$ (coming from the Dishkant hook), gives the D-operators:

$$J_2(a,b) = S(b, T(N(a), N(b))),$$

with which $J_2(N(b), N(a)) = S(N(a), T(a,b)) = J_1(a,b)$ or, equivalently, $J_2(a,b) = J_1(N(b), N(a))$: D-operators are the contrasymmetricals of Q-operators. Hence, it can be repeated all that has been said for J_1. For example,

If $S = \max$, $T = \min$, $N = N_0$, it is $J_2(a,b) = J_1(1-a, 1-b) = \max(b, \min(1-b, 1-a)) = \max(b, 1 - \max(a,b))$, that verifies

$W(a, \max(b, 1-\max(a,b))) = \max(0, \max(a+b-1, a-\max(a,b))) =$
$\left.\begin{cases} W(a,b), & \text{if } b \leqslant a \text{ or } b > a \text{ and } b > \frac{1}{2} \\ 0, & \text{if } b > a \text{ and } b \leqslant \frac{1}{2} \end{cases}\right\} \leqslant b$. It is a W-conditional.

Example 3.2.11 The protoform $\mu \to \sigma = \mu \cdot \sigma$ (coming from the classical conjunctive conditional $a \to b = a \cdot b$), gives

$$J(a,b) = T(a,b),$$

functions with the inconvenience of the property $J(a,b) = J(b,a)$, but verifying,

$$T_0(a, J(a,b)) = T_0(a, T(a,b)) \leqslant T(a,b) \leqslant \min(a,b) \leqslant b$$

that is, all of them are conditionals for any t-norm T_0 and in particular, for the greatest of them. They are always taken as min-conditionals. For example,

- If $T = \min$, $J(a,b) = \min(a,b)$, is called the Mamdani conditional
- If $T = prod_\varphi$, $J(a,b) = \varphi^{-1}(\varphi(a)\cdot\varphi(b))$, are called Larsen conditionals
- $T = W_\varphi$ is never used, since it can be $J(a,b) = 0$ with $a > 0$ and $b > 0$.

For example, with $\varphi(x) = \frac{x(1+x)}{2}$ (an order automorphism), it is $\varphi^{-1}(x) = \frac{\sqrt{8x+1}-1}{2}$, and

$$\begin{aligned} J(a,b) &= \varphi^{-1}(\varphi(a)\cdot\varphi(b)) = \varphi^{-1}(\frac{a(1+a)}{2}.\frac{b(1+b)}{2}) = \varphi^{-1}(\frac{ab(1+a)(1+b)}{4}) \\ &= \frac{\sqrt{ab(1+a)(1+b)+1}-1}{2} \end{aligned}$$

that, of course, is a min-conditional.

Remark 3.2.12 A way of avoiding the undesirable symmetry $J(a,b) = J(b,a)$ in the case of Mamdani-Larsen min-conditionals, is taking

$$J(a,b) = T(a^r, b^s),$$

with real numbers r, s with $1 > s$. Then $T_0(a, T(a^r, b^s)) \leqslant \min(\min(a^r, b^s), b^s) \leqslant b^s \leqslant b$, and J is a min-conditional.

Example 3.2.13 Once given a conditional statement 'If x is P, then y is Q', and represented 'x is P' by $\mu_P(x)$, and 'y is Q' by $\mu_Q(y)$, it remains to be understood what it is meant by the 'statement' $\mu_P(x) \to \mu_Q(y)$. It could be, or not to be, $\mu_P(x) \to \mu_Q(y) = (\mu_P \to \mu_Q)(x,y)$, with $\mu_P \to \mu_Q$ a fuzzy set in $X \times Y$, identified with some expression involving the connectives *and* $(\cdot)$, *or* $(+)$, *not* $(')$. In the affirmative case, it is said that $\mu_P \to \mu_Q$ is expressible in material form, for example, $\mu_P \to \mu_Q = \mu'_P + \mu_Q$, or $\mu_P \to \mu_Q = \mu'_P + \mu_P \cdot \mu_Q$, etc. These material forms are called protoforms; for instance, the protoform of the Kleene-Diennes fuzzy conditional is $a' + b$, and that of Mamdani fuzzy conditional is $a \cdot b$.

If $\mu_P \to \mu_Q$ does not correspond with a protoform, one can try to represent it by means of an R-implication, that is, by $J_{\min}$ or by some J_{prod_φ}, since all the J_{W_φ} do correspond to a protoform $a' + b$, or $\mu'_P + \mu_Q$, with + represented by W^*_φ and $'$ by N_φ.

In addition, there is a problem that should be taken into account when representing $\mu_P \to \mu_Q$. The problem is the following. Suppose that we know $\mu_P \to \mu_Q$ should be represented by a function J that is a min-conditional, but that we are not able to decide a protoform and we take $J_{\min}$. Since $J \leqslant J_{\min}$, we will reach the biggest possible output. This should be known. Analogously, if J should be a prod-conditional, from $J \leqslant J_{prod}$, follows the same comment.

Example 3.2.14 Let's stop for a while in the above mentioned concept of implication function, a concept that comes directly from the properties shown by the Boolean conditional $a \to b = a' + b$, whose truth value is usually represented by

$$v(a \to b) = \max(1 - v(a), v(b)).$$

Look that, if $b_1 \leqslant b_2$, it follows $a' + b_1 \leqslant a' + b_2$, or $a \to b_1 \leqslant a \to b_2$, if $a_1 \leqslant a_2$, is $a'_2 \leqslant a'_1$, and $a_2 \to b \leqslant a_1 \to b$. Since $v(a)$ and $v(b)$ only take the values {0, 1}, $v(a \to b)$ shows the truth-table.

$v(a)$	$v(b)$	$v(a \to b)$
0	0	1
0	1	1
1	0	0
1	1	1

Because of that, a function $J : [0, 1] \times [0, 1] \to [0, 1]$ is called a fuzzy implication function provided:

1. J is decreasing in its first variable, and non-decreasing in its second one
2. $J(0, 0) = J(0, 1) = J(1, 1) = 1, J(1, 0) = 0$

Obviously, S-implications and R-implications are fuzzy implication functions, but Q and D operators are not always so, since, for instance

$$J_S(0.4, 0.3) = \max(1 - 0.4, \min(0.4, 0.3)) = 0.6$$

$$J_D(0.3, 0.1) = \max(0.1, \min(0.7, 0.9)) = 0.7$$

$$J_D(0.3, 0.4) = \max(0.4, \min(0.7, 0.6)) = 0.6,$$

that is, for instance $0.1 < 0.4$ but $J_D(0.3, 0.1) > J_D(0.3, 0.4)$.

Analogously, $J(a,b) = T(a,b)$ are not implication functions, since $J(0,0) = T(0,0) = 0$, and $a_1 \leqslant a_2 \Rightarrow J(a_1,b) \leqslant J(a_2,b)$. Also Sasaki operators are not, for example, $0.4 < 0.8$, but $J_S(0.4, 0.9) = \max(1-0.4, \min(0.4, 0.9)) = \max(0.6, 0.4) = 0.6 < 0.8 = J_S(0.8, 0.9)$, that is, J_S is non-decreasing in the second variable.

To represent conditional statements as they appear in language, the concept of implication function is sometimes excessive. What is needed are just T-conditionals, except in the cases where more properties are necessary to represent the meaning of the conditional statements.

Remark 3.2.15 There is a question that is not independent of the representation J for the conditional statement. The question is: What are we going to do with J? Which is the purpose for using J?

We are to make an inference that, in principle, could be forwards

$$\{\mu, \mu \to \sigma\} \vdash \sigma,$$

or backwards,

$$\{\sigma', \mu \to \sigma\} \vdash \mu'.$$

The first type of inference corresponds to search for the solutions of $\mu \cdot (\mu \to \sigma) \leqslant \sigma$, that is,for J and T_0 such that,

$$T_0(\mu(x), J(\mu(x), \sigma(y)) \leqslant \sigma(y); \quad \forall x, y \in X. \tag{*}$$

The second type corresponds to search for the solutions of $\sigma' \cdot (\mu \to \sigma) \leqslant \mu'$, that is, for J, T_1 and N such that,

$$T_1(N(\sigma(y)), J(\mu(x), \sigma(y)) \leqslant N(\mu(x)); \quad \forall x, y \in X. \tag{**}$$

Hence, given J, we need to know T_0 such that $T_0(a, J(a,b)) \leqslant b$, for forward inference, and given J and N, we need to know T_1 such that $T_1(N(b), J(a,b)) \leqslant N(a)$, for backwards inference.

Notice that the two t-norms in (*) and (**) are not necessarily coincidental. For example, given $J(a,b) = \max(1-a, b)$, with $N = N_0$, can we do backwards inference? To answer this question we just need to know if there is a continuous t-norm T_1 such that $T_1(1-b, \max(0, \max(1-a, b)) \leqslant 1-a$, with $a = 1$ it results $T_1(1-b, b) = 0$, and $T_1 = W$. Then, since,

$W(1-b, \max(1-a,b)) = \max(0, 1-b+\max(1-a,b)-1)$
$= \max(0, \max(1-a-b, 0)) = \left\{\begin{array}{ll} 1-a-b, & \text{if } a+b \leqslant 1 \\ 0, & \text{if } a+b > 1 \end{array}\right\} \leqslant 1-a$, because $b \geqslant 0$ implies $-b \leqslant 0$, and $1-a-b \leqslant 1-a$. Finally, the answer is: Yes, with $T_1 = W$. It can also be done backwards inference in the following cases,

1. With $J_{\min}$, since $W(1-b, J_{\min}(a,b)) = \left\{\begin{array}{ll} 1-b, & \text{if } a \leqslant b \\ W(1-b, b) = 0, & \text{if } a > b \end{array}\right\} \leqslant 1-a.$

2. With J_{prod}, since $W(1-b, J_{prod}(a,b)) = \left\{ \begin{array}{l} 1-b, \text{ if } a \leqslant b \\ b\frac{1-b}{a}, \text{ if } a > b \end{array} \right\} \leqslant 1-a$.
3. With J_W, since $W(1-b, \min(1, 1-a+b)) = \min(1-b, 1-a) \leqslant 1-a$.
4. With $J(a,b) = 1-a+ab$, since $W(1-b, \min(1, 1-a+ab)) = (1-b)(1-a) \leqslant 1-a$.
5. With $J(a,b) = \max(1-a, \min(a,b))$, since $W(1-b, \max(1-a, \min(a,b))) \leqslant W(1-b, \max(1-a, b)) \leqslant 1-a$.
6. With $J(a,b) = prod^*(1-a, a \cdot b) = 1-a+a^2b$, since $W(1-b, prod^*(1-a, a \cdot b)) \leqslant W(1-b, prod^*(1-a, b)) = W(1-b, 1-a+a \cdot b) \leqslant 1-a$.

Nevertheless, the case $J(a,b) = T(a,b)$ is, actually, negative. To have $T_1(1-b, T(a,b)) \leqslant 1-a$, it is necessary (take $a=1$) that $T_1(1-b, b) = 0$, that is, $T_1 = W$, but it is $W(1-b, T(a,b)) = \max(0, T(a,b)-b) = 0$, since $T(a,b) \leqslant b$ implies $T(a,b) - b \leqslant 0$. Hence, by one side from $W(1-b, T(a,b)) = 0 \leqslant 1-a$, it seems that backwards inference is possible. But, given the scheme: "$\sigma', \mu \to \sigma : \mu'$", what results is $\sigma' \cdot (\mu \to \sigma) = \mu_0$, that forces $Conj(\{\sigma', \mu \to \sigma\}) = \emptyset$. In conclusion, Mamdani-Larsen conditionals don't allow backwards inference.

Remark 3.2.16 It should be pointed out that except $J_{\min}$, J_{prod_φ}, and $J(a,b) = T(a,b)$, most of the functions J are T_0-conditionals for $T_0 = W$, and almost all do also verify backwards inference also with $T_0 = W$. And the t-norms in the Łukasiewicz's family show the disturbing problem of having zero-divisors!

Remark 3.2.17 The name R-implication, or residuated implication, comes from the idea of 'residuum' that clearly appear in the case of J_{prod} when

$$\text{If } a > b, \text{ then } J_{prod}(a,b) = \frac{b}{a}.$$

Remark 3.2.18 In the same vein under which it was proven that R-implications J with $T \neq W_\varphi$ are not S-implications, it is easy to show that they are not expressible in material protoform, that is, by an expression with logical connectives. Take the perhaps more general material protoform $\mu' \cdot (\sigma + \sigma') + \mu \cdot \sigma$. Is it possible that

$$J_{T_0}(a,b) = S_1(T_1(N_1(a), S_2(b, N_2(b)), T_2(a,b))),$$

for $T_0 \neq W_\varphi$, S_1 and S_2 continuous t-conorms, T_1, T_2 continuous t-norms, and N_1, N_2 strong negations? With $b=0$, it follows

$$J_{T_0}(a,0) = \sup\{z \in [0,1]; T(z,a) \leqslant 0\} = 0$$

$$S_1(T_1(N_1(a), S_2(0,1)), 0) = T_1(N_1(a), 1) = N_1(a),$$

or, $S_1(N_1(a), 0) = 0$. This means that S_1 is not a t-conorm. Hence, the decision of representing an R-implication can't be taken from a material protoform interpretation of it.

3.3 Short Note on Other Modes of Reasoning

The mode of reasoning given by the scheme

$$\{\mu, \mu \to \sigma\} \vdash \sigma$$

is classically called *Modus Ponendo Ponens* (from the Latin, mode of starting the truth (of σ) by placing the truth (of μ)), or, for short *Modus Ponens*. That given by the scheme $\{\sigma', \mu \to \sigma\} \vdash \mu'$ is classically called *Modus Tollendo Tollens* (from the Latin, mode of stating the falsity (of μ) by placing the falsity (of σ)), or, for short *Modus Tollens*. They correspond to what we called forwards and backwards reasoning. But there are again other *modes* of reasoning that can be considered, for example,

- *Modus Tollendo Ponens*, given by the scheme $\{\mu', \mu+\sigma\} \vdash \sigma$ and also called Mode of Disjunctive Reasoning, and classically proven by $\mu' \cdot (\mu+\sigma) = \mu \cdot \mu' + \mu' \cdot \sigma = \mu' \cdot \sigma \leqslant \mu$ (in a Boolean algebra).
- *Modus Ponendo Tollens*, given by the scheme $\{\mu, (\mu \cdot \sigma)'\} \vdash \sigma'$, classically proven by $\mu \cdot (\mu \cdot \sigma)' = \mu \cdot (\mu' + \sigma') = \mu \cdot \sigma' \leqslant \sigma'$ (in a Boolean algebra).
- *Constructive Dilemma*, given by the scheme $\{\mu + \lambda, \mu \to \sigma, \lambda \to \eta\} \vdash \sigma + \eta$, classically proven by $(\mu+\lambda) \cdot (\mu \to \sigma) \cdot (\lambda \to \eta) = (\mu+\lambda) \cdot (\mu'+\sigma) \cdot (\lambda'+\eta) = \mu\sigma\lambda' + \eta\lambda\mu' + \eta\mu\sigma + \eta\lambda\sigma \leqslant \sigma + \eta$ (in a Boolean algebra where $a \to b = a' + b$).
- *Destructive Dilemma*, given by the scheme $\{\mu' + \sigma', \lambda \to \mu, \eta \to \sigma, \} \vdash \lambda' + \sigma'$, classically proven by $(\mu' + \sigma')(\lambda' + \mu)(\eta' + \sigma) = \lambda'(\mu' \cdot \eta' + \mu'\sigma) + \eta'(\sigma' \cdot \lambda' + \mu \cdot \sigma') \leqslant \lambda' + \eta'$ (in a Boolean algebra where $a \to b = a' + b$).

What in the fuzzy case? For example, in the case of the Disjunctive Mode we need to find all the possibilities for $\mu' \cdot (\mu + \sigma) \leqslant \sigma$, that is, to solve the functional equation

$$T(N(a), S(a, b)) \leqslant b$$

for all a, b in $[0, 1]$, in the three variables T, S, N. With $b = 0$, it follows $T(N(a), a) = 0$, or $T = W_\varphi$, $N \leqslant N_\varphi$. Taking $N = N_\varphi$ it results $W_\varphi(N_\varphi(a), S(a, b)) = \varphi^{-1}(\max(0, 1 - \varphi(a) + \varphi(S(a, b)) - 1))$
$= \varphi^{-1}(\max(0, +\varphi(S(a, b)) - \varphi(a))) \leqslant b$, implying $\varphi(S(a, b)) - \varphi(a) \leqslant \varphi(b)$ or $\varphi(S(a, b)) \leqslant \varphi(a) + \varphi(b)$. Hence, $S(a, b) \leqslant \varphi^{-1}(\min(1, \varphi(a) + \varphi(b))) = W^*(a, b)$. For example, it can be taken $S = W^*_\varphi$ or $S = \max$.

- $W_\varphi(N_\varphi(a), \max(a, b)) = \varphi^{-1}(\max(0, \varphi(b) - \varphi(a))) \leqslant \varphi^{-1}(\varphi(b)) = b$
- $W_\varphi(N_\varphi(a), W^*_\varphi(a, b)) = \varphi^{-1}(\min(1 - \varphi(a), \varphi(b))) \leqslant \varphi^{-1}(\varphi(b)) = b$

Hence, the disjunctive mode can be used in, for example, the cases
$(W_\varphi, N_\varphi, \max)$ and $(W_\varphi, N_\varphi, W^*_\varphi)$.

Remark 3.3.1 It should be pointed out that the Modus Ponendo Tollens (MPT) can be reduced, in the case of duality, to the disjunctive mode by means of the change $\mu = \alpha'$, $\sigma = \beta'$, in which case since $(\mu \cdot \sigma)' = \mu' + \sigma'$ it follows $\mu \cdot (\mu \cdot \sigma)' = \mu \cdot (\mu' + \sigma') = \alpha' \cdot (\alpha + \beta) \leqslant \beta = \sigma'$. Hence, it holds with the triplet $(W_\varphi, N_\varphi, W_\varphi^*)$.

3.4 Inference with Fuzzy Rules

A central topic fuzzy logic deals with are non-rigid, dynamic systems involving 'variables' $x_1, \ldots, x_n, y$ taking values in, respectively, universes $X_1, \ldots, X_n, Y$, and constrained by imprecise rules r_i of the type

If x_1 is P_{1i}, and x_2 is P_{2i},..., and x_n is P_{ni}, then y is Q_i $(1 \leqslant i \leqslant n)$,

with predicates $P_{ji} (1 \leqslant j \leqslant n)$ in X_j, and Q_i in Y.

Let's consider the simplest case with two variables $x \in X$, $y \in Y$, constrained by a single rule *If x is P, then y is Q*.

When observing the system (x, y), the variable x in the rule's antecedent not always will show '*x is P*', but 'x is P^*' with P^* some predicate slightly modificate from P. For example, if $P = short$, it could be $P^* = very\ short$, $almost\ short$, etc. A concrete example is

Rule:	If tomatoes are red, they are ripe
Observation:	Tomatoes are very red
Conclusion:	Tomatoes are very ripe,

where $P = red$, $Q = ripe$, $P^* = very\ red$, and $Q^* = very\ ripe$. Hence, the corresponding scheme of forwards reasoning is

Rule:	If x is P, then y is Q
Observation:	x is P^*
Consequence:	y is Q^*

where P, Q, P^* are known, and Q^* is unknown. This scheme is the *Generalized Modus Ponens* (GMP), and to find Q^* through fuzzy artillery it is needed to have the representations

- $\mu_P \in [0, 1]^X$, of P.
- $\mu_Q \in [0, 1]^Y$, of Q.
- $\mu_{P^*} \in [0, 1]^X$, of P^*.
- $(\mu_P \to \mu_Q)(x, y) = J(\mu_P(x), \mu_Q(y))$, with a convenient T-conditional J, for the rule.

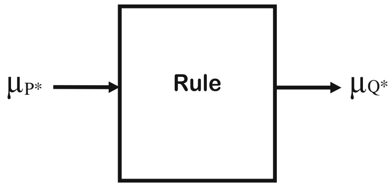
From this representations should follow a representation of Q^*, that is, $\mu_{Q^*} \in [0,1]^Y$, by taking into account that it should be a logical consequence of the set of premises $\{\mu_{P^*}, \mu_P \to \mu_Q\}$ by the *Generalized Modus Ponens* (GMP). That is, μ_{Q^*} does verify

$$0 \neq T_0(\mu_{P^*}(x), J(\mu_P(x), \mu_Q(y))) \leqslant \mu_{Q^*}(y), \tag{3.1}$$

for all $x \in X$, $y \in Y$, and a continuous t-norm T_0 verifying

$$0 \neq T_0(\mu_P(x), J(\mu_P(x), \mu_Q(y))) \leqslant \mu_Q(y),$$

for all $x \in X, y \in Y$, stating that if $P^* = P$, then $Q^* = Q$. This is done by preserving the Modus Ponens (MP) in the occasion in which 'x is P' is observed. The fuzzy set μ_{P^*} is called the *input*



and μ_{Q^*} is the *output*.

Obviously,

$$\mu_{Q^*}(y) = \sup_{x \in X} T_0(\mu_{P^*}(x), J(\mu_P(x), \mu_Q(y))), \ \forall y \in Y,$$

is the greatest function verifying the GMP (3.1). This formula is known as the *Compositional Rule of Fuzzy Inference* (CRI, for short), and was introduced by Lotfi A. Zadeh as the output fuzzy logic considers in the systems that are described by rules. It is not to be forgotten that T_0 is the continuous t-norm that makes J a T_0-conditional.

Sometimes, the input is just numerical, crisp in the form '$x = x_0$', that is, 'x is P^*' is 'x is x_0', or '$x \in \{x_0\}$' and then $\mu_{P^*} = \mu_{\{x_0\}}$, with

$$\mu_{\{x_0\}}(x) = \begin{cases} 1, & \text{if } x = x_0 \\ 0, & \text{if } x \neq x_0. \end{cases}$$

In this case,

$$\mu_{Q^*}(y) = \sup_{x \in X} T_0(\mu_{\{x_0\}}(x), J(\mu_P(x), \mu_Q(y))) = J(\mu_P(x_0), \mu_Q(y)), \ \forall y \in Y,$$

is a simpler expression that does not force to compute $\sup_{x \in X}$.

Sometimes, in addition to the input, the rule's consequent μ_Q is also numerical, that is 'y is Q' is '$y = y_0$', or 'y is y_0' , or '$y \in \{y_0\}$'. In this case

$$\mu_Q(y) = \mu_{\{y_0\}}(y) = \begin{cases} 1, & \text{if } y = y_0 \\ 0, & \text{if } y \neq y_0 \end{cases},$$

and the output is

$$\mu_{Q^*}(y) = J(\mu_P(x_0), \mu_{\{y_0\}}(y)) = \begin{cases} J(\mu_P(x_0), 1), & \text{if } y = y_0 \\ J(\mu_P(x_0), 0), & \text{if } y \neq y_0. \end{cases}$$

Example 3.4.1 Take $X = [0, 1]$, $Y = [0, 10]$, and the rule *'If x is small, then y is big'*, with the observation that *'x is big'* and $J(a, b) = \max(1 - a, b)$. With the membership functions $\mu_P(x) = 1 - x$, $\mu_Q(y) = \frac{y}{10}$, $\mu_{P^*}(x) = x$, it results

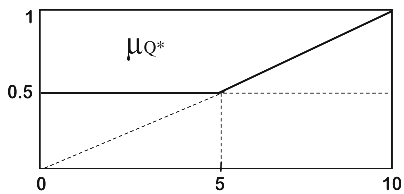
$$\mu_{Q^*}(y) = \sup_{x \in [0,1]} W(x, \max(x, \frac{y}{10})) = W(1, \max(1, \frac{y}{10})) = 1,$$

or, $\mu_{Q^*} = \mu_1$, that means $Q^* = \text{all}$.

Example 3.4.2 With the same rule of last example and the input $x_0 = 0.5$, it is

$$\mu_{Q^*}(y) = \max(0.5, \max(0.5, \frac{y}{10})), \ \forall y \in [0, 1]$$

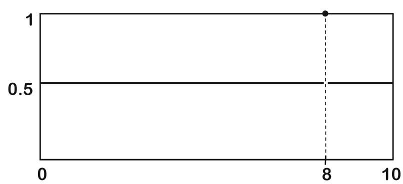
graphically,



Example 3.4.3 With the rule 'If x is small, then $y = 8$', and the input $x = 0.5$, is

$$\mu_{Q^*}(y) = \max(0.5, \mu_{\{8\}}(y)) = \begin{cases} 1, & \text{if } y = 8 \\ 0.5, & \text{if } y \neq 8 \end{cases},$$

graphically,



Remark 3.4.4 In general, it is not easy to assign a name to the functional output μ_{Q^*}, that is, to express Q^* linguistically. In Example 3.4.2, it could be said $Q^* = big$ after 5 and constantly equal to 0.5 before 5. In Example 3.4.3, it could be said $Q^* =$ almost always 0.5.

3.4.1 Finite Case

Let's consider the particular case where both universes X and Y are finite sets. If

$$X = \{x_1, \ldots, x_n\},\ \ Y = \{y_1, \ldots, y_m\},$$

fuzzy sets μ_P, μ_Q and μ_{P^*} are of the forms:

- $\mu_P = r_1/x_1 + \cdots + r_n/x_n$, meaning $\mu_P(x_i) = r_i$, $1 \leqslant i \leqslant n$
- $\mu_Q = s_1/y_1 + \cdots + s_n/y_m$, meaning $\mu_Q(y_j) = s_i$, $1 \leqslant j \leqslant m$
- $\mu_{P^*} = r_1^*/x_1 + \cdots + r_n^*/x_n$, meaning $\mu_{P^*}(x_i) = r_i^*$, $1 \leqslant i \leqslant n$

Provided J is an adequate T_0-conditional to represent the rule 'If x is P, then y is Q' ($(\mu_P \to \mu_Q)(x, y) = J(\mu_P(x), \mu_Q(y))$), it is

$$\begin{aligned}(\mu_P \to \mu_Q)(x_i, y_j) = J(\mu_P(x_i), \mu_Q(y_j)) &= J(r_i, s_j)\\ &= a_{ij}, 1 \leqslant i \leqslant n, 1 \leqslant j \leqslant m.\end{aligned}$$

With all that,

$$\mu_{Q^*}(y_j) = \sup_{1 \leqslant i \leqslant n} T_0(\mu_{P^*}(x_i), J(\mu_P(x_i), \mu_Q(y_j)) = \max_{1 \leqslant i \leqslant n} T_0(r_i^*, J(r_i, s_j))$$

$$= \max_{1 \leqslant i \leqslant n} T_0(r_i^*, a_{ij}), 1 \leqslant i \leqslant n,$$

since in the finite case the sup is just max.

Calling $\mu_{Q^*} = s_1^*/y_1 + \ldots + s_m^*/y_m$, it results

$$s_j^* = \max_{1 \leqslant i \leqslant n} T_0(r_i^*, a_{ij}),\ \ 1 \leqslant j \leqslant m,$$

showing an special 'composition' of the matrices

$$(r_1^*, \ldots, r_n^*) = [\mu_{P^*}] \ \text{ and } \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1m} \\ a_{21} & a_{22} & \ldots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nm} \end{pmatrix} = [J],$$

in which the elements of the classical product of matrices (rows by columns) $\sum_{1 \leqslant i \leqslant n} r_i^* \cdot a_{ij}$ are substituted by $\max_{1 \leqslant i \leqslant n} T_0(r_i^*, a_{ij})$.

This composition is called the max-T_0 *product of matrices*, instead of the classical sum-prod composition. Hence,

$$[\mu_{Q^*}] = (s_1^*, \ldots, s_m^*) = [\mu_{P^*}] \otimes [J],$$

gives the CRI's output.

Example 3.4.5 With $\mu_P = 0.7/x_1 + 0.8/x_2 + 1/x_3$, $\mu_Q = 0.9/y_1 + 0.6/y_2 + 0.8/y_4$, $\mu_{P^*} = 0.6/x_1 + 0.7/x_2 + 1/x_3$, and $J(a, b) = \min(1, 1 - a + b)$, follows:

$a_{11} = J(0.7, 0.9) = \min(1, 1 - 0.7 + 0.9) = 1$; $a_{12} = J(0.7, 0.6) = 0.9$; $a_{13} = J(0.7, 0) = 0.3$; $a_{14} = J(0.7, 0.8) = 1$
$a_{21} = J(0.8, 0.9) = 1$; $a_{22} = J(0.8, 0.6) = 0.8$; $a_{23} = J(0.8, 0) = 0.2$; $a_{24} = J(0.8, 0.8) = 1$
$a_{31} = J(1, 0.9) = 0.9$; $a_{32} = J(1, 0.6) = 0.6$; $a_{33} = J(1, 0) = 0$; $a_{34} = J(1, 0.8) = 0.8$.

Hence,

$$[\mu_{Q^*}] = (0.6\ 0.7\ 1) \otimes \begin{pmatrix} 1 & 0.9 & 0.3 & 1 \\ 1 & 0.8 & 0.2 & 1 \\ 0.9 & 0.6 & 0 & 0.8 \end{pmatrix} = (0.9\ 1\ 0.9\ 0.8),$$

since: $(\max(W(0.6, 1), W(0.7, 1), W(1, 0.9)),\ \max(W(0.6, 0.9), W(0.7, 0.8), W(1, 0.6)),\ \max(W(0.6, 0.3), W(0.7, 0.2), W(1, 0)),\ \max(W(0.6, 1), W(0.7, 1), W(1, 0.8))) = (\max(0.6, 0.7, 0.9),\ \max(1, 1, 0.6),\ \max(0.9, 0.9, 0),\ \max(0.6, 0.7, 0.8)) = (0.9\ 1\ 0.9\ 0.8)$. That is

$$\mu_{Q^*} = 0.9/y_1 + 1/y_2 + 0.9/y_3 + 0.8/y_4.$$

In the case μ_P is interpreted P = *more or less big*, μ_Q is interpreted Q = *not very big*, and μ_{P^*} is interpreted P^*= *medium*, it is possible to agree on Q^* = *more or less big*.

3.4.2 Inference with Several Rules

Actually, there are no systems described by a single rule. What to do when a system is described by, at least, two rules? With, for example

- r1: If x is P_1, then y is Q_1
- r2: If x is P_2, then y is Q_2,

an input μ_{P^*} gives

- with r1, the output $\mu_{Q_1^*}$, obtained by using CRI
- with r2, the output $\mu_{Q_2^*}$, obtained by using CRI.

Provided the firing of the rules corresponds to 'fire r1 **or** fire r2', then the final output is given by

$$\mu_{Q^*} = \max(\mu_{Q_1^*}, \mu_{Q_2^*}),$$

and analogously for more than two rules. For example, in the case of p rules $r_1, \ldots, r_p$, the result will be

$$\mu_{Q^*} = \max(\mu_{Q_1^*}, \ldots, \mu_{Q_p^*}),$$

where $\mu_{Q_i^*}(1 \leqslant i \leqslant p)$ is the output obtained with the rule r_i and the input μ_{P^*}.

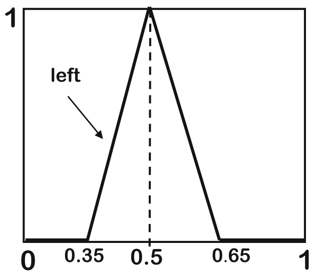
Example 3.4.6 With $X = [0, 1]$, $Y = [0, 10]$, consider the rules

- If x is big, then $y = 2$
- If x is $small$, then $y = 8$
- If x is $around$ 0.5, then $y = 6$,

and the input $x_0 = 0.4$. Which is the final output of this system if the rules are represented by $J(a, b) = a \cdot b$?

- $\mu_{Q_1^*}(y) = J(\mu_B(0.4), \mu_{\{2\}}(y)) = 0.4.\mu_{\{2\}}(y) = \begin{cases} 0.4, & \text{if } y = 2 \\ 0, & \text{if } y \neq 2 \end{cases}$
- $\mu_{Q_2^*}(y) = J(\mu_s(0.4), \mu_{\{8\}}(y)) = 0.6.\mu_{\{2\}}(y) = \begin{cases} 0.6, & \text{if } y = 8 \\ 0, & \text{if } y \neq 8 \end{cases}$
- $\mu_{Q_3^*}(y) = J(\mu_{A0.5}(0.4), \mu_{\{6\}}(y)) = \mu_{A0.5}(0.4).\mu_{\{6\}}(y)$
  $= \begin{cases} \mu_{A0.5}(0.4), & \text{if } y = 6 \\ 0, & \text{if } y \neq 6, \end{cases}$

with $\mu_B(x) = x$, and $\mu_S(x) = 1 - x$. Taking as $\mu_{A0.5}$ the triangular function



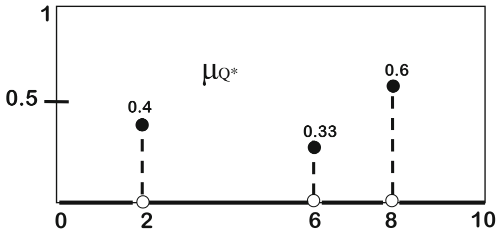
and since the left side equation is $y = \frac{x - 0.35}{0.15}$, it is $\mu_{A0.5}(0.4) = 0.\overset{\frown}{3}$. Hence,

$$\mu_{Q_3^*}(y) = \begin{cases} 0.\overset{\frown}{3}, & \text{if } y = 6 \\ 0, & \text{if } y \neq 6, \end{cases}$$

Finally

$$\mu_{Q^*}(y) = \max(\mu_{Q_1^*}(y), \mu_{Q_2^*}(y), \mu_{Q_3^*}(y)) = \begin{cases} 0.4, & \text{if } y = 2 \\ 0.\overset{\frown}{3}, & \text{if } y = 6 \\ 0.6, & \text{if } y = 8 \\ 0, & \text{otherwise} \end{cases}$$

the output is given by the graphics,



In many applications, the output should be converted into a single numerical value: it should be 'defuzzified'. In this cases with numerical input and numerical rule's consequents (the most used in *fuzzy control*), such number is easyly obtained by averaging the values of μ_{Q^*}, in the form

$$\frac{2 \times 0.4 + 6 \times 0.\overset{\frown}{3} + 8 \times 0.6}{0.4 + 0.\overset{\frown}{3} + 0.6} = \frac{7.5998}{1.3333} = 5.6999 \simeq 5.7$$

Hence, the numerical output that corresponds to the input $x_0 = 0.4$, is $y_0 = 5.7$.

Remark 3.4.7 Notice that once a system of rules linguistically describing the behavior of a system is given, and where the consequents of the rules are numerical, at each numerical input x_0 in X does correspond a numerical output y_0 in Y. In that way, a function CRI: $X \to Y$ is defined. As it will be later on commented, were the system's behaviour previously known by a continuous function $f : X \to Y$, the function CRI approaches, under some additional conditions, the function f.

Remark 3.4.8 Look how important is to properly select the T-conditionals representing the rules.
Given the rule 'If x is small, then y is big', with $X = Y = [0, 1]$, and $\mu_S(x) = 1 - x$, $\mu_B(y) = y$, $J(a, b) = \max(1 - a, b)$,
it follows $J(\mu_S(x), \mu_B(y)) = \max(x, y)$, that could be interpreted as 'x is big *or* y is big'.
With $J(a, b) = \min(1, 1 - a + b)$, it follows $J(\mu_S(x), \mu_B(y)) = \min(1, x + y) = W^*(x, y)$, also interpretable as 'x is big *or* y is big'. But with $J(a, b) = \min(a, b)$, is $J(\mu_S(x), \mu_B(y)) = \min(1 - x, y)$, interpreted as *'x is small and y is big'*.

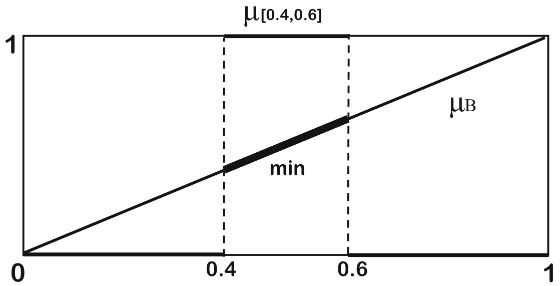
3.4.3 Examples

Example 3.4.9 Rule 'If x is big, then $y = 0.8$', with x, y in [0, 1], and the observation $x \in [0.4, 0.6]$. Hence:

$$J(\mu_B(x), \mu_{\{0.8\}}(y)) = x \cdot \mu_{\{0.8\}}(y) = \begin{cases} x, & \text{if } y = 0.8 \\ 0, & \text{if } y \neq 0.8 \end{cases}$$

Then, $\mu_{Q^*}(y) = \underset{x \in [0,1]}{Sup} \min(\mu_{[0.4,0.6]}(x), x\mu_{\{0.8\}}(y)) = \begin{cases} 0.6, & \text{if } y = 0.8 \\ 0, & \text{if } y \neq 0.8, \end{cases}$

since,



Example 3.4.10 Rule: 'If x is big, then y is small', with the same observation as that in the last example and with $\mu_B(x) = x$, $\mu_S(y) = 1 - y$, and $J(a, b) = \min(a, b)$, follows:
$\mu_{Q^*}(y) = \underset{x \in [0,1]}{Sup} \min(\mu_{[0.4,0.6]}(x), \min(x, 1 - y)) = \underset{x \in [0,1]}{Sup} \min(\min(\mu_{[0.4,0.6]}(x), x), 1 - y) = \underset{x \in [0,1]}{Sup} \min(x, 1 - y) = \min(0.6, 1 - y)$.

Example 3.4.11 $X = \{1, 2, 3\}$, $Y = \{6, 7\}$. Rule: 'If x is around 2, then $y = 6$', and $\mu_{P^*}(x) = 0.6/1 + 0.9/2 + 0.7/3$, with $J(a, b) = ab$ (Larsen). It results $(\mu_{Q^*}(6), \mu_{Q^*}(7)) = (0.6\ 0.9\ 0.7) \otimes \begin{pmatrix} 0.5 & 0 \\ 1 & 0 \\ 0.5 & 0 \end{pmatrix} = (0.9\ 0.7)$, that is $\mu_{Q^*} = 0.9/6 + 0/7$.

Example 3.4.12 With $x \in [0, 1]$, and $y \in [0, 1]$, consider

- r1:'If x is big, then y is small', represented by $J_1(a, b) = ab$
- r2:'If x is very small, then y is very big', represented by $J_2(a, b) = \min(a, b)$

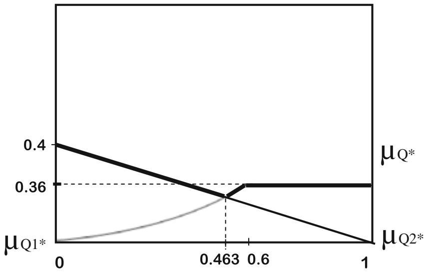
Consider $x_0 = 0.4$, $\mu_B(x) = x$, $\mu_S(y) = 1 - y$, $\mu_{VS}(x) = (1 - x)^2$, $\mu_{VB}(y) = y^2$. Then

- $J_1(\mu_B(x), \mu_S(y)) = x(1-y)$
- $J_2(\mu_{VS}(x), \mu_{VB}(y)) = \min((1-x)^2, y^2) = [\min(1-x, y)]^2$.

Hence,

- $\mu_{Q_1^*}(x) = J_1(0.4, 1-y) = 0.4 \cdot (1-y)$,
- $\mu_{Q_2^*}(x) = J_2((1-0.4)^2, y^2) = (\min(0.6, y))^2 = \begin{cases} 0.36, & \text{if } y \geqslant 0.6 \\ y^2, & \text{if } y < 0.6, \end{cases}$

whose graphics are,



Finally,

$$\mu_{Q^*}(y) = \max(\mu_{Q_1^*}(y), \mu_{Q_2^*}(y)) = \begin{cases} 0.4(1-y), & \text{if } 0 \leqslant y \leqslant 0.463 \\ y^2, & \text{if } 0.463 \leqslant y \leqslant 0.6 \\ 0.36, & \text{if } 0.6 \leqslant y \leqslant 1. \end{cases}$$

Example 3.4.13 Let's find the function CRI:$X \to Y$, in the case with $X = [0, 1]$, $Y = [0, 1]$, and

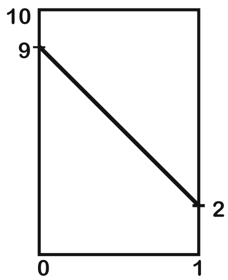
- r1: If x is small, then $y = 9$
- r2: If x is big, then $y = 2$,

it follows $\mu_{Q_1^*}(y) = (1-x)\mu_{\{9\}}(y) = \begin{cases} 1-x, & \text{if } y = 9 \\ 0, & \text{if } y \neq 9, \end{cases}$

$\mu_{Q_2^*}(y) = x\mu_{\{2\}}(y) = \begin{cases} x, & \text{if } y = 2 \\ 0, & \text{if } y \neq 2, \end{cases}$
and

$$\mu_{Q^*}(y) = \max(\mu_{Q_1^*}(y), \mu_{Q_2^*}(y)) = \begin{cases} x, & \text{if } y = 2 \\ 1-x, & \text{if } y = 9 \\ 0, & \text{otherwise} \end{cases} \quad \text{that gives,}$$

$$CRI(x) = \frac{2x + 9(1-x)}{x + 1 - x} = 9 - 7x,$$



as the "theoretical" (linear) behavior of the system (x, y). For each $x_0 \in X$, the value $CRI(x_0) \in Y$ is the defuzzified value that corresponds to x_0.

Remark 3.4.14 Systems of fuzzy rules behave as *universal approximators*. This means the following. Suppose a system (x, y), with $x \in [a, b]$, $y \in [c, d]$, that behave by following the continuous function $f(x) = y$. For each $\varepsilon > 0$,there is always a system of fuzzy rules and a defuzzification method for the output, giving a function CRI:$X \to Y$ such that

$$|f(x) - CRI(x)| < \varepsilon, \text{ for all } x \in [a, b]$$

This theorem (whose proof is here avoided) is simply an existential one, since there is no general method for obtaing neither a fuzzy representation of the system of rules, not the defuzzifiction method. It simply shows that it is possible to find a CRI approaching enough well f for all points in $[a, b]$.

3.5 Deffuzification

How to defuzzify non discrete outputs μ_{Q^*}? Let us proceed with two examples without computational difficulties.

1st Example. Rules,

- r1: If x is big, then y is small
- r2: If x is small, then y is big

with $X = [0, 1]$, and $Y = [0, 10]$. Take,

$$\mu_B(x) = x, \quad \mu_S(y) = 1 - \frac{y}{10}, \quad \mu_S(x) = 1 - x, \quad \mu_B(y) = \frac{y}{10},$$

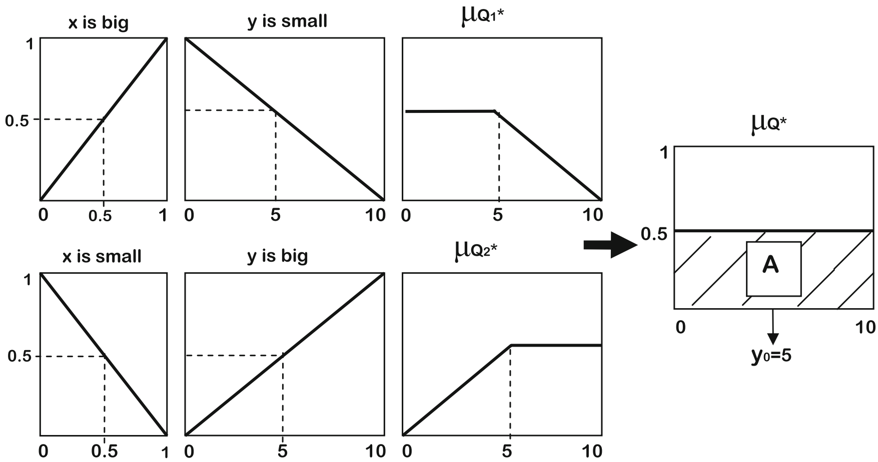
and $J(a, b) = \min(a, b)$ -Mamdani-. Notice that, with the observation $x_0 = 0.5$,

$$\mu_{Q_1^*}(y) = \min(0.5, 1 - \frac{y}{10}), \quad \mu_{Q_2^*}(y) = \min(1 - 0.5, \frac{y}{10}).$$

Then, $\mu_{Q^*}(y) = \max(\min(0.5, 1 - \frac{y}{10}), \min(0.5, \frac{y}{10})) = \max(0.5, \min(1 - \frac{y}{10}, \frac{y}{10})) = 0.5$, since $\min(1 - \frac{y}{10}, \frac{y}{10}) \leqslant 0.5$.
The area below $\mu_{Q^*}(y) = 0.5$, is $A = 0.5 \times 10 = 5$ square units. Hence, a way to *defuzzify* μ_{Q^*} *consists of searching the center of area*, that is, a point $y_0 \in [0, 10]$ such that

$$\int_0^{y_0} 0.5dy = \frac{A}{2} = 2.5, \text{ or } \int_0^{y_0} dy = 5.$$

Hence, $[y]_0^{y_0} = y_0 = 5$. The defuzzified value corresponding to $x_0 = 0.5$, is $y_0 = 5$. The method, when the conditional is Mamdani, is graphically reflected as follows.



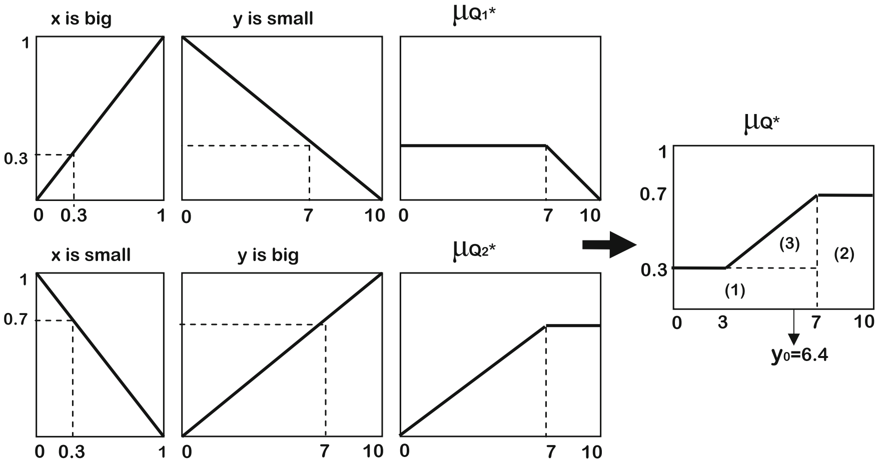
2nd Example. Identical to the first example, but with the input $x_0 = 0.3$. It is

$$\mu_{Q_1^*}(y) = \min(0.3, 1 - \frac{y}{10}), \quad \mu_{Q_2^*}(y) = \min(1 - 0.3, \frac{y}{10}).$$

Hence,

$$\mu_{Q^*}(y) = \begin{cases} 0.3, & \text{if } 0 \leqslant y \leqslant 3 \\ \frac{y}{10}, & \text{if } 3 \leqslant y \leqslant 7 \\ 0.7, & \text{if } 7 \leqslant y \leqslant 10 \end{cases}$$

that is graphically find as follows:



The area below μ_{Q^*} is A = rectangle(1) + rectangle(2) + triangle (3) = $0.3 \times 7 + 3 \times 0.7 + \frac{4 \times 0.4}{2} = 5$. Since the area of the rectangle with base [0, 3] is $0.3 \times 3 = 0.9$, it is $y_0 > 3$. Hence,

$$\int_0^3 0.3dy + \int_3^{y_0} \frac{y}{10}dy = A/2 = 2.5.$$

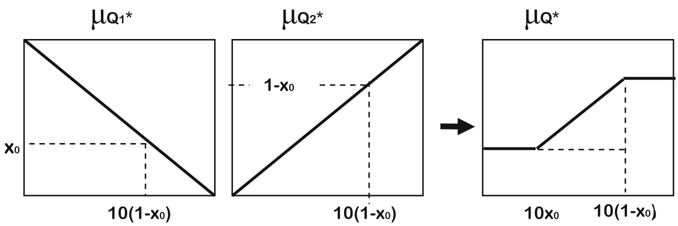
Then $3 \times 0.3 + \frac{1}{10}\int_3^{y_0} ydy = 2.5$, or $\int_3^{y_0} ydy = 10(2.5 - 0.9) = 16$. Thus,

$$[\frac{y^2}{2}]_3^{y_0} = 16 \Rightarrow y_0^2 - 9 = 32 \Rightarrow y_0^2 = 41 \Rightarrow y_0 = 6.4.$$

The defuzzified value that corresponds to $x_0 = 0.3$, is $y_0 = 6.4$.

Defuzzifying with the centre of the area we obtained an output for all values x_0 in [0, 1]. Let's see it by means of the function CRI with defuzzification made by the centre of the area.

1. The graphics, for any input $x_0 \leqslant 1/2$, is



The area below μ_{Q^*} is

$$A = 10(1 - x_0)x_0 + (10 - 10(1 - x_0))(1 - x_0) + \frac{(1 - x_0 - x_0)(10(1 - x_0) - 10x_0)}{2}$$

then,

$$2A = 20(1-x_0)x_0 + 20(1-x_0)x_0 + (1-2x_0)(10-20x_0),$$

and

$$A = 20x_0(1-x_0) + 5 - 20x_0 + 20x_0^2 = 5.$$

Hence,

$$10x_0^2 + \int_{10x_0}^{y_0} \frac{y}{10} dy = 2.5 \Rightarrow y_0 = \sqrt{50 - 100x_0^2},$$

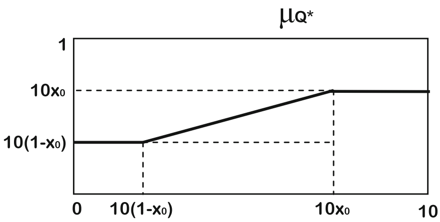
and

$$CRI(x) = \sqrt{50 - 100x^2}, \text{ if } x \leqslant 1/2.$$

For example, $CRI(0.5) = 0.5$, and $CRI(0.3) = \sqrt{41} = 6.4$, as it was shown. It is also $CRI(0.1) = 7$.

Last formula, $CRI(x) = \sqrt{20 - 100x^2}$, gives real values provided $20 - 100x^2 \geqslant 0$, or $x^2 \leqslant 1/5$. Since, it is $x \leqslant 1/5$, that implies $x^2 \leqslant 1/4 \leqslant 1/2$, and it follows that the formula is useful for all $x \in [0, 1]$ such that $x \leqslant 1/2$.

2. For any input $x_0 \geqslant 1/2$, the graphic is



and the area below μ_{Q^*} is

$$A = 10x_0(1-x_0) + (10-10x_0)x_0 + \frac{(10x_0 - 10(1-x_0))(x_0 - 1 + x_0)}{2}$$
$$= 20x_0(1-x_0) + (10x_0 - 5)(2x_0 - 1) = 5.$$

Hence, $A/2 = 2.5$, and

$$10(1-x_0)^2 + \int_{10(1-x_0)}^{y_0} \frac{y}{10} dy = 2.5, \text{ or } 100(1-x_0)^2 + \int_{10(1-x_0)}^{y_0} y dy = 2.5,$$

giving $y_0^2 = 50 - 100(1 - x_0)$, or $y_0 = \sqrt{50 - 100(1 - x_0)^2}$, that gives real values provided $50 - 100(1 - x_0)^2 \geqslant 0$, equivalent to $1/2 \geqslant (1 - x_0)^2$, or to $x_0 \geqslant 1 - \sqrt{1/2}$. Since, $1/2 \leqslant x$, it is $(1 - x)^2 \leqslant 1/4 \leqslant 1/2$. Then,

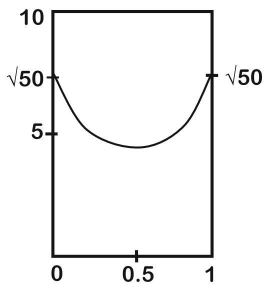
$$CRI(x) = \sqrt{50 - 100(1 - x)^2}, \text{ if } 1/2 \leqslant x.$$

For example,
$CRI(0.7) = \sqrt{41} = 6.4$, $CRI(0.8) = \sqrt{46} = 6.78$, $CRI(0.9) = \sqrt{49} = 7$.

3. Finally, with defuzzification by the centre of area, is:

$$CRI(x) = \begin{cases} \sqrt{50(1 - 2x^2)}, & \text{if } x \leqslant 1/2 \\ \sqrt{50(1 - 2(1 - x)^2)}, & \text{if } x \geqslant 1/2. \end{cases}$$

Notice that $CRI(0) = \sqrt{50}$, $CRI(0.5) = \sqrt{50(1 - 20.5^2)} = \sqrt{25} = 5$, and $CRI(1) = \sqrt{50}$.
The graphic of CRI is



3.6 Rules and Conjectures

As it was said before, the output is a logical consequence of the premises given by the input and the rule. Notwithstanding, the situation is different if, taking the rule as defining the system, only the input is considered as a premise. But, before to consider this question, let us consider what happens when there is more than a single rule.

1. If $\mu_{Q_1^*}, \mu_{Q_2^*} \in Cons(\{\mu_{P^*}\})$, from $\mu_{P^*} \leqslant \max(\mu_{Q_1^*}, \mu_{Q_2^*}) = \mu_{Q^*}$, follows $\mu_{Q^*} \in Cons(\{\mu_{P^*}\})$.
2. If $\mu_{Q_1^*}$ or $\mu_{Q_2^*}$ is a conjecture of $\{\mu_{P^*}\}$, then $\mu_{Q^*} \in Conj(\{\mu_{P^*}\})$. The proof follows in this way, provided it is, for instance, $\mu_{Q_1^*} \in Conj(\{\mu_{P^*}\})$,

- $\mu_{Q^*} = \max(\mu_{Q_1^*}, \mu_{Q_2^*}) \Rightarrow \mu_{Q_1^*} \leqslant \mu_{Q^*},\ \mu_{Q_2^*} \leqslant \mu_{Q^*} \Rightarrow \mu'_{Q^*} \leqslant \mu'_{Q_1^*},\ \mu'_{Q^*} \leqslant \mu'_{Q_2^*} \Rightarrow \mu'_{Q^*} \leqslant \min(\mu'_{Q_1^*}, \mu'_{Q_2^*})$.
- If $\mu_{P^*} \leqslant \mu'_{Q^*}$, then $\mu_{P^*} \leqslant \mu'_{Q_1^*}$, that is absurd. Hence, it is $\mu_{P^*} \nleqslant \mu'_{Q^*}$, and $\mu_{Q^*} \in Conj(\{\mu_{P^*}\})$.

In conclusion

- If all the partial outputs $\mu_{Q_1^*}, \mu_{Q_2^*}, ..., \mu_{Q_n^*}$, are consequences of the input μ_{P^*}, also the final output μ_{Q^*} is a consequence of μ_{P^*}.
- If at least one of the partial outputs $\mu_{Q_i^*}$ $(1 \leqslant i \leqslant n)$ is just a conjecture of the input μ_{P^*}, also the final output μ_{Q^*} is a conjecture of μ_{P^*}.

Nevertheless, it is not usual that μ_{Q^*} results to be a consequence of the single input μ_{P^*}. Let's us introduce a necessary and sufficient condition for it in the particular case in which there is only one rule represented by $J(a, b) = ab$ (Larsen).

Let it be "If x is P, then y is Q" $(x, y \in X)$, and $(\mu_P \to \mu_Q)(x, y) = \mu_P(x)\mu_Q(y)$, with the input $x = x_0$. Then,

$$\mu_{Q^*}(y) = \mu_{P^*}(x_0)\mu_Q(y), \quad \forall y \in X.$$

Provided $\mu_{\{x_0\}} \neq \mu_0$, to have $\mu_{\{x_0\}}(y) \leqslant \mu_{Q^*}(y) = \mu_{P^*}(x_0)\mu_Q(y)$, it is necessary that, with $y = x_0$, $1 \leqslant \mu_{P^*}(x_0)\mu_{Q^*}(x_0)$ or $1 = \mu_{P^*}(x_0) = \mu_{Q^*}(x_0)$. Hence, $\mu_{Q^*} \in Cons(\{\mu_{\{x_0\}}\})$ implies $\mu_{P^*}(x_0) = \mu_{Q^*}(x_0) = 1$.

Provided $\mu_{P^*}(x_0) = \mu_{Q^*}(x_0) = 1$, from $\mu_{Q^*}(y) = \mu_{P^*}(x_0)\mu_Q(y)$, follows $\mu_{Q^*}(y) = \mu_Q(y)$, for all $y \in X$, and,

- If $y = x_0$, $\mu_{\{x_0\}}(x_0) = 1 = \mu_{Q^*}(x_0)$
- If $y \neq x_0$, $\mu_{\{x_0\}}(y) = 0 \leqslant \mu_{Q^*}(y)$,

that is

$$\mu_{\{x_0\}}(y) \leqslant \mu_{Q^*}(y), \text{ for all } y \in X.$$

Hence, in this particular case, the necessary and sufficient condition for being $\mu_{Q^*} \in Cons(\{\mu_{\{x_0\}}\})$ is that $\mu_{P^*}(x_0) = \mu_{Q^*}(x_0) = 1$. Nevertheless, what happens in most of the cases is that $\mu_{Q^*} \in Conj(\{\mu_{\{x_0\}}\})$, with $\mu_{Q^*} \in Sp(\{\mu_{\{x_0\}}\})$, or $\mu_{Q^*} \in Hyp(\{\mu_{\{x_0\}}\})$.

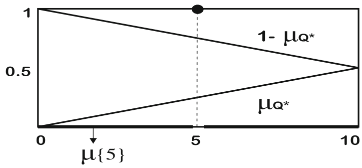
3.7 Two Final Examples

Let's show an example in which the output is a speculation of the input and other in which the output is a hypothesis.

Example. Rule, "If x is small, then y is big", with $X = Y = [0, 10]$ and $J(a, b) = ab$ (Larsen), with $\mu_S(x) = 1 - \frac{x}{10}$, $\mu_B(y) = \frac{y}{10}$, and $x_0 = 5$. It is

$$\mu_{Q^*}(y) = J(1 - \frac{5}{10}, \frac{y}{10}) = \frac{y}{20},$$

and from the graphic



it is clear that μ_{Q^*} is not comparable with $\mu_{\{5\}}$ (μ_{Q^*} NC $\mu_{\{5\}}$), and that $\mu_{\{5\}} \nleqslant \mu'_{Q^*} = 1 - \mu_{Q^*}$. Hence,

$$\mu_{Q^*} \in Conj(\{\mu_{\{5\}}\}), \text{ and namely } \mu_{Q^*} \in Sp(\{\mu_{\{5\}}\}).$$

Notice that $\mu_{Q^*}(5) = \frac{1}{4} \neq 1$.

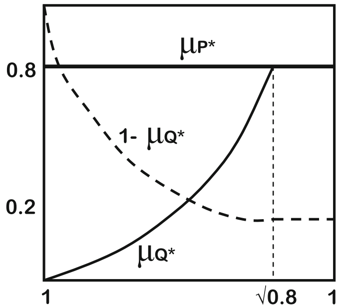
Example. Rule, "If x is big, then y is very big", with $X = Y = [0, 10]$ and the observation "x is constantly equal to 0.8" for all $x \in [0, 1]$. Taking

$$\mu_B(x) = x, \ \mu_{vB}(y) = y^2, \ J(a, b) = \min(a, b) \text{ (Mamdani)},$$

follows

$$\mu_{Q^*}(y) = \min(\mu_{P^*}(0.8), \mu_{vB}(y)) = \min(0.8, y^2).$$

Graphically,



since $y^2 = 0.8$ means $y = \sqrt{0.8}$. Hence, it is $\mu_0 \neq \mu_{Q^*} \leqslant \mu_{P^*}$, and $\mu_{P^*} = \mu'_{Q^*} = 1 - \mu_{Q^*}$, that imply $\mu_{Q^*} \in Conj(\{\mu_{P^*}\})$, and namely $\mu_{Q^*} \in Hyp(\{\mu_{P^*}\})$.

Notice that this second example contains the observation that the input is a constant.

Last Remark

For any continuous t-norm T, the function $\mu : Y \to [0, 1]$, defined by

$$\mu(y) = \underset{x \in X}{Sup} T(\mu_{P^*}(x), J(\mu_P(x), \mu_Q(y))), \ \forall y \in Y,$$

does verify

$$T(\mu_{P^*}(x), J(\mu_P(x), \mu_Q(y))) \leqslant \mu(y), \ \forall y \in Y, x \in X,$$

that is, $\mu \in Cons(\{\mu_{P^*}, \mu_P \to \mu_Q\})$. Nevertheless, if $T = T_0$, the continuous t-norm for which J is a T_0-conditional, that is, such that

$$T_0(\mu_P(x), J(\mu_P(x), \mu_Q(y))) \leqslant \mu(y) =, \ \forall y \in Y, x \in X,$$

it could be that when $\mu_{P^*} = \mu_P$, then $\mu \neq \mu_Q$. A undesiderable situation, because fuzzy logic must contain all classical cases.

For example, with the rule 'If x is small, then y is big' ($X = Y = [0, 1]$), and $J(a, b) = \max(1 - a, b)$ that is a W-conditional, taking $\mu_S(x) = 1 - x$ and $\mu_B(y) = y$, follows:

- With $T = W$, $\mu(y) = \underset{x \in [0,1]}{Sup}\ W((1 - x), \max(x, y)) = \underset{x \in [0,1]}{Sup}\ (0, y - x) = y = \mu_B(y)$.
- With $T = prod$, $\mu(y) = \underset{x \in [0,1]}{Sup}\ (1 - x) \max(x, y) = \underset{x \in [0,1]}{Sup}\ \max((1 - x)x, (1 - x)y) = y = \max(1/4, y/2)$, not coincidental with μ_B.
- with $T = \min$, $\mu(y) = \underset{x \in [0,1]}{Sup}\ \min((1 - x), \max(x, y)) = 1$, or $\mu = \mu_1$, also not coincidental with μ_B.

Hence, although with any continuous t-norm T, an output μ is obtained, if this T does not make J a T-conditional it is not sure that $P = P^*$ implies $Q = Q^*$. For this reason, it is necessary to take T_0 with CRI!

Chapter 4
Fuzzy Relations

4.1 What Is a Fuzzy Relation?

A predicate R on a cartesian product $X_1 \times \cdots \times X_n$ is called a relational (n-ary) predicate. For example, if $X_1 = X_2 = [0, 10]$, $R =$ close to, '(x, y) is R', or 'x is close to y', is a relational binary predicate.

Analogously, if $X_1 = X_2 = London$, R = lives in the same borough, or 'x lives in the same borough than y', is a relational binary predicate.

A fuzzy relation in $X_1 \times \cdots \times X_n$ is any function $\mu : X_1 \times \cdots \times X_n \to [0, 1]$. If interpreting $\mu_R(x_1, \ldots, x_n) =$'degree up to which $(x_1, \ldots, x_n)$ is in $\underset{\backsim}{R}$', it is said that μ_R represents the n-ary relational relation R'.

Any rule 'If x is P, then y is Q' defines the binary predicate Q/P in $X \times Y$ given by

$$(x, y) \text{ is } Q/P \ \Leftrightarrow \text{If } x \text{ is } P, \text{ then } y \text{ is } Q,$$

whose representation, or membership function of the corresponding fuzzy set $\underset{\backsim}{Q/P}$ is given by

$$\mu_{Q/P}(x, y) = (\mu_P \to \mu_Q)(x, y) = J(\mu_P(x), \mu_Q(y)),$$

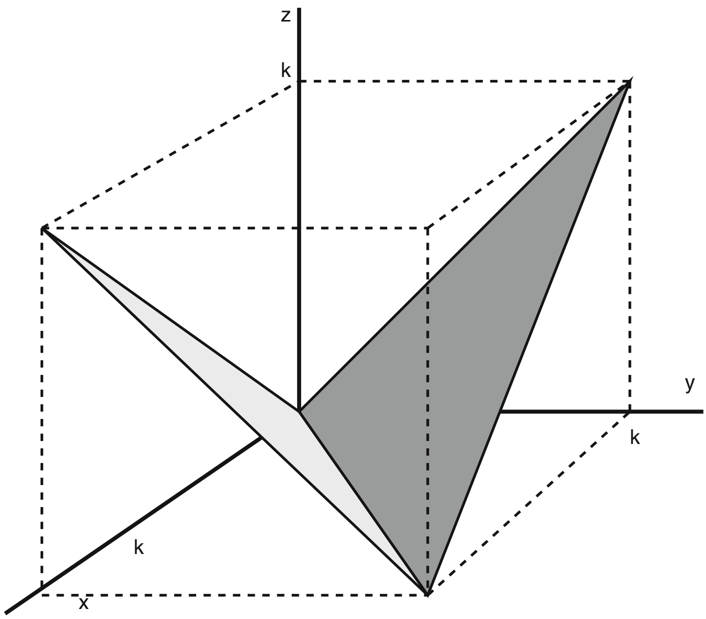
once a T-conditional J adapted to the meaning of Q/P is selected. *A fuzzy relation μ_R is nothing else than a fuzzy set $\underset{\backsim}{R}$ in $X_1 \times \cdots \times X_n$.*

For example, if $R =$ *'close to'*, is represented by

$$\mu_R(x, y) = \max(0, k|x - y|), \quad \text{for all } x, y \in [0, 1],$$

E. Trillas and L. Eciolaza, *Fuzzy Logic*, Studies in Fuzziness and Soft Computing 320, DOI 10.1007/978-3-319-14203-6_4

with $k \in (0, 1)$ a parameter chosen at each particular case, it is $\mu_R(0, 0) = 0$, $\mu_R(1, 1) = 0$, $\mu_R(0, 1) = \mu_R(1, 0) = \max(0, k) = k$, $\mu_R(1/2, 1) = \mu_R(1, 1/2) = \max(0, k/2) = k/2$, etc., with the graphic,



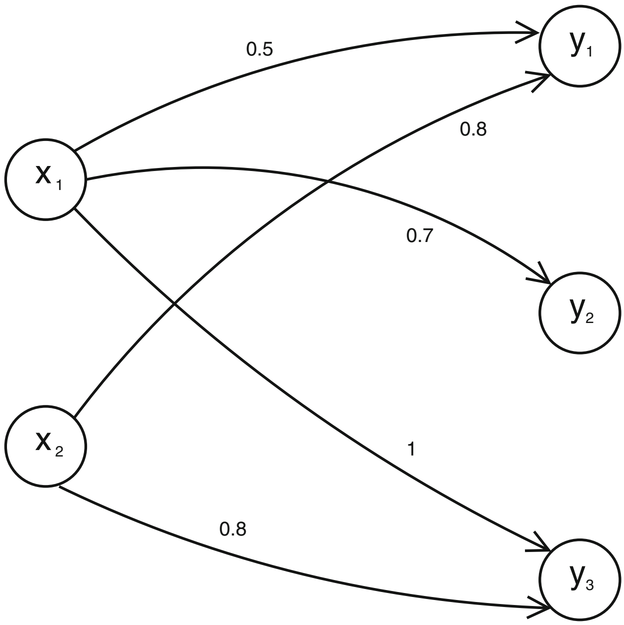
When the sets $X_1, \ldots, X_n$ are finite, μ_R is reduced to a matrix. For example if $X_1 = \{x_1, \ldots, x_p\}$, and $X_2 = \{y_1, \ldots, y_q\}$, then

$$\mu_R(x_i, y_j) = r_{ij},\ 1 \leqslant i \leqslant n,\ 1 \leqslant j \leqslant m, \text{ or,}$$
$$\mu_R = r_{11}/(x_1, y_1) + \cdots + r_{nm}/(x_n, y_m),$$

that gives the $n \times m$ matrix

$$[R] = \begin{pmatrix} r_{11} & r_{12} & \ldots & r_{1m} \\ r_{21} & r_{22} & \ldots & r_{2m} \\ \vdots & \vdots & & \vdots \\ r_{n1} & r_{n2} & \ldots & r_{nm} \end{pmatrix}.$$

In the finite case there is again another representation of a fuzzy relation by means of a directed graph. For example, if $X_1 = \{x_1, x_2\}$ and $X_2 = \{y_1, y_2, y_3\}$, the fuzzy relation $[R] = \begin{pmatrix} 0.5 & 0.7 & 1 \\ 0.8 & 0 & 0.8 \end{pmatrix}$, corresponds to the directed graph



4.2 How to Compose Fuzzy Relations?

Given two fuzzy relations $\mu : X \times Y \to [0, 1]$, and $\sigma : Y \times Z \to [0, 1]$, how can we obtain a relation $\lambda : X \times Z \to [0, 1]$ through μ and σ? To solve this problem, there is the $Sup - T$ product of fuzzy relations, given by

$$\lambda(x, z) = \underset{y \in Y}{Sup}\, T(\mu(x, y), \sigma(y, z)), \quad \text{for all } (x, z) \in X \times Z,$$

a formula that, in the finite case $X = \{x_1, \dots, x_n\}$, $Y = \{y_1, \dots, y_m\}$, $Z = \{z_1, \dots, z_p\}$, reduces to,

$$\lambda(x_i, z_j) = \underset{1 \leqslant k \leqslant m}{Max}\, T(\mu(x_i, y_k), \sigma(y_k, z_j)).$$

Provided $[\mu] = (r_{ik})$, $[\sigma] = (s_{kj})$, then

$$t_{ij} = \lambda(x_i, z_j) = \underset{1 \leqslant k \leqslant m}{Max}\, T(r_{i,k}, s_{kj}), 1 \leqslant i \leqslant n, 1 \leqslant j \leqslant p,$$

giving the matrix $[\lambda] = (t_{ij})$ as the Max-T product, or composition, of the matrices $[r_{ik}]$ and $[s_{kj}]$, that was introduced before by: $(t_{ij}) = (r_{ik}) \otimes_T (s_{kj})$.

Example 4.2.1 [1]Let

- $X = \{p_1, \ldots, p_4\}$, a set of patients
- $Y = \{s_1, s_2, s_3\}$, a set of symptoms
- $Z = \{d_1, \ldots, d_5\}$, a set of deceases,

and the fuzzy relation σ

$$[\sigma] = \begin{pmatrix} 0.7 & 0 & 0 & 0.3 & 0.6 \\ 0.5 & 0.5 & 0.8 & 0.4 & 0 \\ 0 & 0.7 & 0.2 & 0.9 & 0 \end{pmatrix}$$

showing the medical knowledge of how strongly each symptom is associated with a decease. Suppose also that, by examining the patients, the doctors conclude the matrix

$$[\mu] = \begin{pmatrix} 0 & 0.3 & 0.4 \\ 0.2 & 0.5 & 0.3 \\ 0.8 & 0 & 0 \\ 0.7 & 0.7 & 0.9 \end{pmatrix}$$

that describes numerically how strongly the symptoms are manifested in the patients. Then,

$$[\lambda] = [\mu] \otimes_{\min} [\sigma]$$

is the matrix expressing the association patients/deceases, and facilitates a medical diagnose. That is,

$$[\lambda] = \begin{pmatrix} 0 & 0.3 & 0.4 \\ 0.2 & 0.5 & 0.3 \\ 0.8 & 0 & 0 \\ 0.7 & 0.7 & 0.9 \end{pmatrix} \otimes_{\min} \begin{pmatrix} 0.7 & 0 & 0 & 0.3 & 0.6 \\ 0.5 & 0.5 & 0.8 & 0.4 & 0 \\ 0 & 0.7 & 0.2 & 0.9 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0.3 & 0.4 & 0.3 & 0.4 & 0 \\ 0.5 & 0.5 & 0.5 & 0.4 & 0.2 \\ 0.7 & 0 & 0 & 0.3 & 0.6 \\ 0.7 & 0.7 & 0.7 & 0.9 & 0.6 \end{pmatrix},$$

where, for instance,

$$t_{43} = \max(\min(0.7, 0), \min(0.7, 0.8), \min(0.9, 0.2)) = \max(0, 0.7, 0.2) = 0.7.$$

The matrix $[\lambda]$ results from a mixing between knowledge and observation.

[1] From, Fuzzy Set Theory, by G.J. Klir, U.H. St. Clair, B. Yuan, Prentice/Hall, N.J., 1997.

Remark 4.2.2 1. As it is easy to prove, it is

$$([\mu] \otimes_T [\sigma])^t = [\sigma]^t \otimes_T [\mu]^t,$$

with the matrices $[\sigma]^t$, $[\mu]^t$, defined by $\sigma^t(x, y) = \sigma(y, x)$, $\mu^t(x, y) = \mu(y, x)$, giving

$$([\mu]^t)^t = [\mu].$$

The matrix $[\mu]^t$ is the *transposed* of $[\mu]$.

2. In general, the max-T composition is associative, but not commutative. That is, *if the compositions* $[\mu] \otimes_T ([\sigma] \otimes_T [\lambda])$, and $([\mu] \otimes_T [\sigma]) \otimes_T [\lambda]$, *are possible* it is,

$$([\mu] \otimes_T [\sigma]) \otimes_T [\lambda] = [\mu] \otimes_T ([\sigma] \otimes_T [\lambda]),$$

but, in general, $[\mu] \otimes_T [\sigma] \neq [\sigma] \otimes_T [\mu]$.

4.3 Which Relevant Properties Do Have a Fuzzy Binary Relation?

The most relevant properties of a fuzzy relation $\mu : X \times X \to [0, 1]$ are the following,

1. Reflexive property, $\mu(x, x) = 1$, for all $x \in X$.
2. Symmetric property, $\mu(x, y) = \mu(y, x)$, for all $x, y \in X$, implies $x = y$.
3. Antisymmetric property, $\mu(x, y) > 0$, $\mu(y, x) > 0$ implies $x = y$.
4. T-transitive property $T(\mu(x, y), \mu(y, z)) \leqslant \mu(x, z)$, for all $x, y, z \in X$, and some continuous t-norm T.

In the finite case, for what concerns properties reflexive and symmetric, the matrix $[\mu] = (t_{ij})$ shows the respective properties,

1′. It is $t_{ii} = 1$, for all $1 \leqslant i \leqslant n$, that is, the main diagonal of $[\mu]$ is constituted by n numbers equal to 1.

2′. It is $t_{ij} = t_{ji}$, for all $1 \leqslant i, j \leqslant n$, that is, the elements of $[\mu]$ are placed symmetrically with respect to the main diagonal.

For example, the matrix

$$\begin{pmatrix} 1 & 0.7 \\ 0.6 & 1 \end{pmatrix} \text{ is reflexive, but not symmetric,}$$

and the matrix

$$\begin{pmatrix} 1 & 0.6 & 0.7 \\ 0.6 & 0 & 0.9 \\ 0.7 & 0.9 & 0.5 \end{pmatrix} \text{ is symmetric, but not reflexive.}$$

For what concerns the antisymmetric property, $t_{ij} > 0$ and $t_{ji} > 0$, implies $i = j$. For example, the matrix

$$\begin{pmatrix} 1 & 0 & 0.7 & 0 \\ 0.6 & 1 & 0 & 0.7 \\ 0 & 0.5 & 1 & 0.8 \\ 0.7 & 0 & 0 & 1 \end{pmatrix}$$

is antisymmetric.

Let's define the binary relation, $[\mu] \leqslant [\sigma]$, between $n \times n$ matrices if $t_{ij} \leqslant s_{ij}$ for all $1 \leqslant i, j \leqslant n$, provided $[\mu] = (t_{ij})$, $[\sigma] = (s_{ij})$. With such definition,

- $[\mu]$ reflects a T-transitive fuzzy relation μ, *if and only* if, $[\mu] \otimes_T [\mu] \leqslant [\mu]$.

  The proof is as follows.

a. If $[\mu]$ is T-transitive, from

$$T(\mu(x_i, x_j), \mu(x_j, x_k)) \leqslant \mu(x_i, x_k),$$

or $T(t_{ij}, t_{jk}) \leqslant t_{ik}$, it is $\underset{1 \leqslant j \leqslant n}{Max}\, T(t_{ij}, t_{jk}) \leqslant t_{ik}$. That is, $[\mu] \otimes_T [\mu] \leqslant [\mu]$.

b. If $[\mu] \otimes_T [\mu] \leqslant [\mu]$, or $\underset{1 \leqslant j \leqslant n}{Max}\, T(t_{ij}, t_{jk}) \leqslant t_{ik}$, follows $T(t_{ij}, t_{jk}) \leqslant t_{ik}$, for all $1 \leqslant i, j \leqslant n$. That is, μ is T-transitive.

- If μ is reflexive and T-transitive, it is $[\mu] \otimes_T [\mu] = [\mu]$. since

$$t_{ik} = T(1, t_{ik}) = T(t_{ii}, t_{ik}) \leqslant \underset{1 \leqslant j \leqslant n}{Max}\, T(t_{ij}, t_{jk}) \leqslant t_{ik},$$

  implies

$$t_{ik} = \underset{1 \leqslant j \leqslant n}{Max}\, T(t_{ij}, t_{jk}), \text{ or } [\mu] = [\mu] \otimes_T [\mu].$$

Remark 4.3.1 The definitions given in this section contain the case of the corresponding classical crisp definitions,

- If $R \subset X \times X$ is a classical reflexive relation in X, its membership function μ_R reflects $(x, x) \in R$ for all $x \in X$, by $\mu_R(x, x) = 1$.
- If $R \subset X \times X$ is a classical symmetric relation in X,

$$(x, y) \in R \Leftrightarrow (y, x) \in R \text{ is reflected by } \mu(x, y) = \mu(y, x).$$

- If $R \subset X \times X$ is antisymmetric,

$$(x, y) \in R \;\&\; (y, x) \in R \Leftrightarrow x = y$$

  is reflected by $\mu_R(x, y) = \mu_R(y, x) = 1 (>0) \;\Rightarrow\; x = y$.

- If R is transitive, $(x, y) \in R$ & $(y, z) \in R \Rightarrow (x, z) \in R$, is reflected by $\mu_R(x, y) = \mu_R(y, z) = 1 \Rightarrow \mu_R(x, z) = 1$, that implies,

$$T(\mu_R(x, y), \mu_R(y, z)) \leqslant \mu_R(x, z),$$

for all t-norms T. Notice that if $\mu_R(x, y) = 0$ or $\mu_R(y, z) = 0$, then, for example,

$$T(\mu_R(x, y), \mu_R(y, z)) = T(0, \mu_R(y, z)) = 0 \leqslant \mu_R(x, z).$$

Hence, for all x, y, z in X, and any continuous t-norm T, is

$$T(\mu_R(x, y), \mu_R(y, z)) \leqslant \mu_R(x, z).$$

that reflects equationally the transitivity of R.

Example 4.3.2 1. The matrix

$$[\mu] = \begin{pmatrix} 1 & 1/8 & 2/8 \\ 1/8 & 1 & 3/8 \\ 2/8 & 3/8 & 1 \end{pmatrix}$$

is reflexive and symmetric (*fuzzy similarity*). In addition,

$$[\mu] \otimes_{prod} [\mu] = [\mu],$$

that is, $[\mu]$ is prod-transitive. Notice that,

$$[\mu] \otimes_{min} [\mu] = \begin{pmatrix} 1 & 1/8 & 2/8 \\ 1/8 & 1 & 3/8 \\ 2/8 & 2/8 & 1 \end{pmatrix} \neq [\mu],$$

and $[\mu]$ is not min-transitive. Of course, since $W \leqslant prod$, $[\mu]$ is also W-transitive. Notice that this last matrix is reflexive, non symmetric, but min-transitive, since

$$\begin{pmatrix} 1 & 1/8 & 2/8 \\ 1/8 & 1 & 3/8 \\ 2/8 & 2/8 & 1 \end{pmatrix} \otimes_{\min} \begin{pmatrix} 1 & 1/8 & 2/8 \\ 1/8 & 1 & 3/8 \\ 2/8 & 2/8 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2/8 & 2/8 \\ 2/8 & 1 & 3/8 \\ 2/8 & 2/8 & 1 \end{pmatrix},$$

hence, this 3×3-matrix reflects a min-*preorder* and, consequently, a T-preorder for all t-norm T.

2. The before mentioned fuzzy relation $\mu(x, y) = \max(0, 1 - k|x - y|)$ is not only reflexive and symmetric, but also W-transitive, as it can be proved by distinguishing the four cases: $\frac{1}{k} \geq |x-y|$, $\frac{1}{k} \geq |y-z|$; $\frac{1}{k} < |x-y|$, $\frac{1}{k} < |y-z|$; $\frac{1}{k} \geq |x-y|$, $\frac{1}{k} < |y-z|$; and $\frac{1}{k} < |x-y|$, $\frac{1}{k} \geq |y-z|$.

Remark 4.3.3 If the fuzzy relation μ_R is T-transitive, and T_1 is a t-norm such that $T_1 \leqslant T$, then μ_R is also T_1-transitive, since

$$T_1(\mu_R(x, y), \mu_R(y, z)) \leqslant T(\mu_R(x, y), \mu_R(y, z)) \leqslant \mu_R(x, z),$$

for all x, y, z.

4.4 The Concept of T-State

Given a fuzzy relation $\mu : X \times X \to [0, 1]$ and a continuous t-norm T, a fuzzy set $\sigma : X \to [0, 1]$ is a T-state of μ, if

$$T(\sigma(x), \mu(x, y)) \leqslant \sigma(y), \ \forall (x, y) \in X \times X.$$

All constant fuzzy sets $\mu_k(x) = k$ for all $x \in X$, and $k \in [0, 1]$, are T-states of any fuzzy relation $\mu : X \times X \to [0, 1]$: $T(\mu_k(x), \mu(x, y)) \leqslant \mu_k(x) = k = \mu_k(y)$. For example, $\mu_0 = \mu_\varnothing$ and $\mu_1 = \mu_X$, are always T-states. Hence, the set $T(\mu)$ of all T-states σ of μ is never empty. From now on, in general we will only refer to non constant T-states σ.

Given a fuzzy relation $\mu : X \times Y \to [0, 1]$, once $y \in Y$ is fixed, we can define the fuzzy set $\mu_y : X \times X \to [0, 1]$, defined by,

$$\mu_y(x) = \mu(x, y), \quad \text{for all } x \in X.$$

When X, Y are finite sets, μ_y is the y-column of the matrix $[\mu]$.

If $\mu : X \times X \to [0, 1]$ is a symmetric and T-transitive fuzzy relation, from

$$T(\mu(x, y), \mu(y, z)) \leqslant \mu(x, z), \quad \text{for all } x, y, z \text{ in } X,$$

follows $T(\mu(y, x), \mu(y, z)) \leqslant \mu(z, x)$, or

$$T(\mu_x(y), \mu(y, z)) \leqslant \mu_x(z),$$

that is, μ_x is a T-state of μ.

For example, with $X = \{x_1, x_2, x_3\}$ and $\mu : X \times X \to [0, 1]$ given by

$$[\mu] = \begin{pmatrix} 1 & 1/8 & 2/8 \\ 1/8 & 1 & 3/8 \\ 2/8 & 3/8 & 1 \end{pmatrix}$$

It is $\mu_{x_1} = 1/x_1 + 1/8/x_2 + 2/8/x_3$, and it results,

$$\mu_{x_1}(x_1)\mu(x_1, y) = \mu(x_1, y) = \mu_{x_1}(y).$$

It is also $\mu_{x_2} = 1/8/x_1 + 1/x_2 + 2/8/x_3$, and

$$\mu_{x_2}(x_1)\mu(x_1, y) = \frac{1}{8}\mu(x_1, y),$$

showing,

- $\frac{1}{8}\mu(x_1, x_1) = \frac{1}{8} = \mu_{x_2}(x_1)$
- $\frac{1}{8}\mu(x_1, x_2) = \frac{1}{8^2} < \frac{1}{8} = \mu_{x_2}(x_2)$
- $\frac{1}{8}\mu(x_1, x_3) = \frac{2}{8^2} < \frac{3}{8} = \mu_{x_2}(x_3)$,

etc. That is, the three fuzzy sets $\mu_{x_1}, \mu_{x_2}, \mu_{x_3}$ are $prod$-states of μ.

Remark 4.4.1 When the fuzzy relation μ represents a conditional statement Q/P (a fuzzy rule 'If x is P, then y is Q'), the T-states of μ are among the fuzzy sets verifying the *Modus Ponens* with respect to the continuous t-norm T.

4.5 Fuzzy relations and α-cuts

Given a fuzzy relation $\mu : X \times X \to [0, 1]$, the α-cuts of μ are the classical (crisp) relations $\mu_{(\alpha)}$ defined by,

$$\mu_{(\alpha)} = \{(x, y) \in X \times X;\ \mu(x, y) \geq \alpha\},$$

for all $\alpha \in [0, 1]$. Obviously, $\mu_{(0)} = X \times X$, and if $\alpha_1 \geq \alpha_2$, it is $\mu_{(\alpha_1)} \leqslant \mu_{(\alpha_2)}$.

- μ is symmetric, if and only if all its α-cuts are symmetric (classical) relations.

$$\mu(x, y) = \mu(y, x) \Leftrightarrow (x, y) \in \mu_{(\alpha)} \text{ and } (y, x) \in \mu_{(\alpha)}.$$

- μ is reflexive, if and only if all α-cuts are reflexive (classical) relations.

$$\mu(x, x) = 1 \Leftrightarrow (x, x) \in \mu_{(\alpha)}, \text{ since } \alpha \leqslant 1.$$

- If μ is antisymmetric, all the α-cuts are antisymmetric (crisp) relations,

$$\text{If } \mu(x, y) \geq \alpha > 0, \text{ and } \mu(y, x) \geq \alpha > 0, \text{ it is } x = y.$$

- If μ is a T-transitive fuzzy relation,

$$(x, y) \in \mu_{(\alpha)}, \text{ and } (y, z) \in \mu_{(\alpha)}, \text{ implies } (x, z) \in \mu_{(T(\alpha,\alpha))},$$

since,

$$\mu(x, y) \geq \alpha,\ \mu(y, z) \geq \alpha \Rightarrow T(\alpha, \alpha) \leqslant T(\mu(x, y)\mu(y, z)) \leqslant \mu(x, z).$$

Hence, *only if* μ *is min-transitive*, it is sure that all α-cuts are classical preorders, and μ results decomposed in the family of preorders $\{\mu_\alpha; \alpha \in (0, 1)\}$.

Example 4.5.1 With $X = \{1, 2, 3, 4\}$, the matrix

$$[\mu] = \begin{pmatrix} 1 & 0.6 & 1 & 0.6 \\ 0.3 & 1 & 0.3 & 0.3 \\ 1 & 0.6 & 1 & 0.6 \\ 0.4 & 0.8 & 0.4 & 1 \end{pmatrix},$$

is obviously reflexive but not symmetric, and verifies $[\mu] \otimes_{\min} [\mu] = [\mu]$. Hence, μ is a min-preorder. Its different α-cuts are

$$[\mu_{(1)}] = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},\ [\mu_{(0.8)}] = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix},\ [\mu_{(0.6)}] = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix},$$

$$[\mu_{(0.4)}] = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix},\ [\mu_{(0.3)}] = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix},$$

that give the classical $\leqslant_{(\alpha)}$ preorders, that follows:

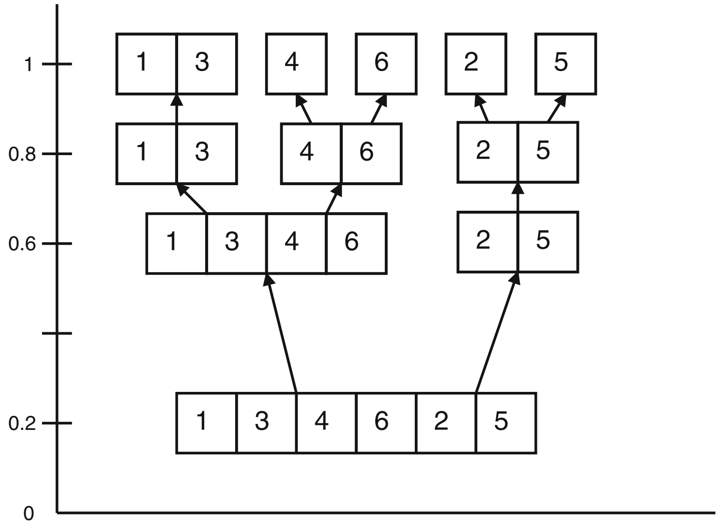
- $\leqslant_{(1)}$: $1 \leqslant_{(1)} 1,\ 2 \leqslant_{(1)} 2,\ 3 \leqslant_{(1)} 3,\ 4 \leqslant_{(1)} 4,\ 1 \leqslant_{(1)} 3,\ 3 \leqslant_{(1)} 1.$
- $\leqslant_{(0.8)}$: $1 \leqslant_{(0.8)} 1,\ 2 \leqslant_{(0.8)} 2,\ 3 \leqslant_{(0.8)} 3,\ 4 \leqslant_{(0.8)} 4,\ 1 \leqslant_{(0.8)} 3,\ 3 \leqslant_{(0.8)} 1,\ 4 \leqslant_{(0.8)} 2.$
- $\leqslant_{(0.6)}$: $1 \leqslant_{(0.6)} 1, \ldots,\ 4 \leqslant_{(0.6)} 4,\ 1 \leqslant_{(0.6)} 3,\ 3 \leqslant_{(0.6)} 1,\ 4 \leqslant_{(0.6)} 2,\ 1 \leqslant_{(0.6)} 2,\ 1 \leqslant_{(0.6)} 4,\ 3 \leqslant_{(0.6)} 2,\ 3 \leqslant_{(0.6)} 4.$
- $\leqslant_{(0.4)}$: $1 \leqslant_{(0.4)} 1, \ldots,\ 4 \leqslant_{(0.4)} 4,\ 1 \leqslant_{(0.4)} 3,\ \ldots,\ 4 \leqslant_{(0.4)} 1,\ 4 \leqslant_{(0.4)} 3.$
- $\leqslant_{(0.3)}$: $1 \leqslant_{(0.3)} 1, \ldots,\ 4 \leqslant_{(0.3)} 3,\ 2 \leqslant_{(0.3)} 1,\ \ldots,\ 4 \leqslant_{(0.3)} 2,$

and, since, $0.3 \leqslant 0.4 \leqslant 0.6 \leqslant 0.8 \leqslant 1$, verify $\leqslant_{(1)} \subset \leqslant_{(0.8)} \subset \leqslant_{(0.6)} \subset \leqslant_{(0.4)} \subset \leqslant_{(0.3)}$.

Example 4.5.2 The fuzzy relation $\mu : X \times X \to [0, 1]$, with $X = \{1, 2, 3, 4, 5, 6\}$, given by

$$[\mu] = \begin{pmatrix} 1 & 0.2 & 1 & 0.6 & 0.2 & 0.6 \\ 0.2 & 1 & 0.2 & 0.2 & 0.8 & 0.2 \\ 1 & 0.2 & 1 & 0.6 & 0.2 & 0.6 \\ 0.6 & 0.2 & 0.6 & 1 & 0.2 & 0.8 \\ 0.2 & 0.8 & 0.2 & 0.2 & 1 & 0.8 \\ 0.6 & 0.2 & 0.6 & 0.8 & 0.8 & 1 \end{pmatrix}$$

is reflexive, symmetrical and min-transitive, since $[\mu] \otimes_{\min} [\mu] = [\mu]$. Hence, all the α-cuts are classical equivalence relations, each one defining a partition of X. The different values of α are 0.2, 0.6, 0.8 and 1 (*levels of crispness* of μ), and it is easy to see that the corresponding partitions $\pi_{0.2}$, $\pi_{0.6}$, $\pi_{0.8}$, and π_1, can be located as the partition tree:

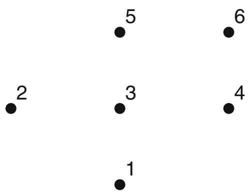


This tree is called the *fuzzy quotient* of X by μ.

Example 4.5.3 The fuzzy relation $\mu : \{1, 2, \dots, 6\} \to [0, 1]$ given by

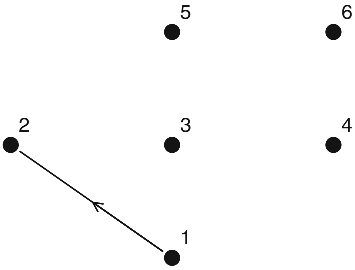
$$[\mu] = \begin{pmatrix} 1 & 0.8 & 0.2 & 0.6 & 0.6 & 0.4 \\ 0 & 1 & 0 & 0 & 0.6 & 0 \\ 0 & 0 & 1 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 1 & 0.6 & 0.4 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

is reflexive, antisymmetric and min-transitive. Hence, is a fuzzy ordering whose α-cuts should be crisp partial orderings. Namely,

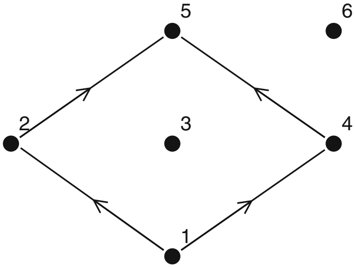
1. $[\mu_{(1)}] = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$, with partial order 

   that only connects the pairs (1, 1), (2, 2), ... (6, 6), and is called the disconnected partial order.

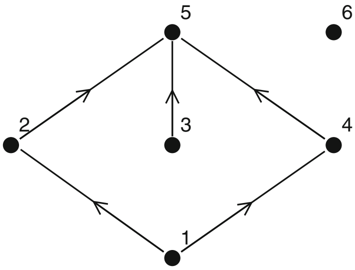
2. $$[\mu_{(0.8)}] = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$



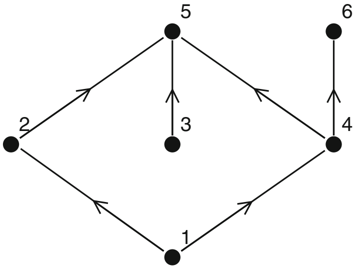
3. $$[\mu_{(0.6)}] = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$



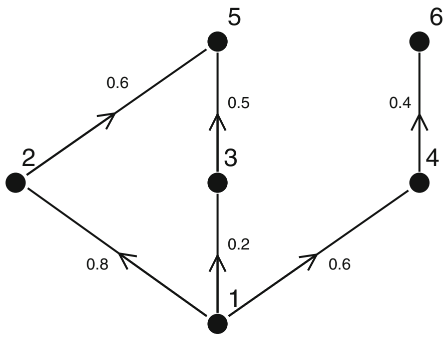
4. $$[\mu_{(0.5)}] = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$



5. $$[\mu_{(0.4)}] = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$



All this crisp partial orderings come from the fuzzy ordering given by $[\mu]$, and which valued directed graph is



where some arrows are avoided because of the min-transitivity of μ. For instance, the degree between 1 and 5, is

$$\mu(1,5) = \underset{1 \leqslant k \leqslant 6}{Max}\, min(\mu(1,k), \mu(k,5)) = max(min(0.8, 0.6),$$
$$min(0.2, 0.5)) = 0.6,$$

and the strengh of the order between 1 and 6, is

$$\mu(1,6) = \underset{1 \leqslant k \leqslant 6}{Max}\, min(\mu(1,k), \mu(k,6)) = min(0.6, 0.4) = 0.4.$$

Of course, $\mu(1,1) = \mu(2,2) = \ldots = \mu(6,6) = 1$.

Chapter 5
T-Preorders and T-Indistinguishabilities

5.1 Which Is the Aim of This Section?

To characterize the T-Preorder and the T-indistinguishability relations by means of particular classes of them, and showing ways of constructing T-preorders and T-indistinguishabilities.

Given a fuzzy relation on X, $\mu : X \times X \to [0, 1]$, with the three properties

1. Reflexivity, $\mu(x, x) = 1$, for all $x \in X$
2. Symmetry, $\mu(x, y) = \mu(y, x)$, for all $(x, y) \in X \times X$
3. T-transitivity, $T(\mu(x, y), \mu(y, z)) \leq \mu(x, z)$, for all x, y, z in X, and a continuous t-norm T,

it can be named

- T-Preorders, those μ verifying 1 and 3.
- T-Indistinguishabilities, those μ verifying 1, 2, and 3.
- Similarities, those μ verifying 1 and 2.

A particular class of T-Preorders is given by the known operators (R-implications)

$$J_T(a, b) = Sup\{z \in [0, 1]; T(z, a) \leq b\},$$

and the corresponding class of T-indistinguishabilities is given by the operators

$$E_T(a, b) = Min(J_T(a, b), J_T(b, a)),$$

shortly written $E_T = \min(J_T, J_T^{\delta})$, with $J_T^{\delta}(a, b) = J_T(b, a)$. It is obvious that

$$J_T(a, a) = 1, \text{ and that } J_T(a, b) \neq J_T(b, a),$$

as well as that $E_T(a, a) = 1$, and $E_T(a, b) = E_T(b, a)$. What it is not so obvious is that relations J_T are T-Transitive:

E. Trillas and L. Eciolaza, *Fuzzy Logic*, Studies in Fuzziness and Soft Computing 320, DOI 10.1007/978-3-319-14203-6_5

$$T(J_T(a,b), J_T(b,c)) \leq J_T(a,c) \text{ for all } a, b, c \text{ in } [0,1]$$

To avoid some difficult technicalities, we will exemplify this general result in the particular case $J_W(a,b) = \min(1, 1-a+b)$:

$$\begin{aligned}
W(J_W(a,b), J_W(b,c)) &= W(\min(1, 1-a+b), \min(1, 1-b+c)) \\
&= \max(0, \min(0, b-a) + \min(1, 1-b+c)) \\
&= \left\{\begin{array}{ll} \min(1, 1-b+c) \leq (1, 1-a+c), & \text{if } a > b \\ \max(0, b-a+\min(1, 1-b+c)) \leq \min(1, 1-a+c), & \text{if } a \leq b \end{array}\right\} \\
&\leq J_W(a,c).
\end{aligned}$$

Then, with the T-transitivity of J_T, it is

$$\begin{aligned}
T(E_T(a,b), E_T(b,c)) &= T(\min(J_T(a,b), J_T(b,a)), \min(J_T(b,c), J_T(c,b))) \\
&\leq T(J_T(a,b), J(b,c)) = E_T(a,c),
\end{aligned}$$

hence, all relations E_T are T-transitive. Then,

- All relations J_T are T-Preorders, and
- All relations E_T are T-Indistinguishabilities

- *If* $\{Ri; i \in I\}$ *is a collection of T-Preorders,* $Inf_{i \in I} R_i(x,y) = R(x,y)$, *is also a T-Preorder.*
  Obviously, $R(x,x) = Inf_{i \in I} R_i(x,x) = 1$, and

$$\begin{aligned}
T(R(a,b), R(b,c)) &= T(\underset{i \in I}{Inf}\, R_i(a,b), \underset{i \in I}{Inf}\, R_i(b,c)) \\
&\leq \underset{i \in I}{Inf}\, T(R_i(a,b), R_i(b,c)) \\
&\leq \underset{i \in I}{Inf}\, R_i(a,c) = R(a,c),
\end{aligned}$$

  since T is a *continuous* t-norm. □
- *If* $\{E_i; i \in I\}$ *is a collection of T-indistinguishabilities,* $Inf_{i \in I} E_i(x,y) = E(x,y)$ *is also a T-indistinguishability.*
  Obviously,

$$E(a,a) = \underset{i \in I}{Inf}\, E_i(a,a) = 1,$$

  and

$$E(a,b) = \underset{i \in I}{Inf}\, E_i(a,b) = \underset{i \in I}{Inf}\, E_i(b,a) = E(b,a).$$

Finally,

$$\begin{aligned}
T(E(a,b), E(b,c)) &= T(\mathop{Inf}_{i\in I} E_i(a,b), \mathop{Inf}_{i\in I} E_i(b,c)) \\
&\le \mathop{Inf}_{i\in I} T(E_i(a,b), E_i(b,c)) \\
&\le \mathop{Inf}_{i\in I} E_i(a,c) = E(a,c)
\end{aligned}$$

since T is a continuous t-norm. □

5.2 The Characterization of T-Preorders

If $\sigma \in [0,1]^X$, define $J_T^\sigma(x,y) = J_T(\sigma(x), \sigma(y))$. Obviously J_T^σ is a T-Preorder in $X \times X$, and if $\mathfrak{F} \subset [0,1]^X$ it is also

$$\mathop{Inf}_{\sigma\in\mathfrak{F}} J_T^\sigma$$

a T-Preorder.

Given a fuzzy relation $\mu : X \times X \to [0,1]$, consider the set $T(\mu)$ of its T-states, that is the fuzzy sets $\sigma : X \to [0,1]$, such that

$$T(\sigma(x), \mu(x,y)) \le \sigma(y), \text{ for all } x, y \in X.$$

As it is known this last inequality is equivalent to

$$\mu(x,y) \le J_T(\sigma(x), \sigma(y)) = J_T^\sigma(x,y),$$

thus,

$$\sigma \in T(\mu) \Leftrightarrow \mu \le J_T^\sigma$$

Hence,

$$\sigma \in T(\mu) \Leftrightarrow \mu \le \mathop{Inf}_{\sigma\in T(\mu)} J_T^\sigma,$$

and it is clear that if $\mu = Inf_{\sigma\in T(\mu)} J_T^\sigma$, μ is a T-Preorder.

Avoiding some technical difficulties in the proof of the converse statement, let us state

- *μ is a T-Preorder if and only if* $\mu = Inf_{\sigma\in T(\mu)} J_T^\sigma$,

a result characterizing the structure of all T-preorders.

For example,

- $T = \min$, all min-preorders are of the form,

$$\mu(x, y) = \underset{\sigma \in T(\mu)}{Inf} \; J^{\sigma}_{\min} = \underset{\sigma \in T(\mu)}{Inf} \begin{cases} 1, & \text{if } \sigma(x) \leq \sigma(y) \\ \sigma(y), & \text{if } \sigma(x) > \sigma(y) \end{cases}$$

- $T = prod_{\varphi}$, all $prod_{\varphi}$-preorders are of the form,

$$\mu(x, y) = \underset{\sigma \in T(\mu)}{Inf} \; J^{\sigma}_{prod_{\varphi}} = \underset{\sigma \in T(\mu)}{Inf} \begin{cases} 1, & \text{if } \sigma(x) \leq \sigma(y) \\ \varphi^{-1}(\frac{\varphi(y)}{\varphi(x)}), & \text{if } \sigma(x) > \sigma(y) \end{cases}$$

- $T = W_{\varphi}$, all W_{φ}-preorders are of the form

$$\mu(x, y) = \underset{\sigma \in T(\mu)}{Inf} \; J^{\sigma}_{W}(x, y) = \underset{\sigma \in T(\mu)}{Inf} \; min(1, 1 - \sigma(x) + \sigma(y))$$

Remark 5.2.1 As an immediate consequence of all that has been said, for any family of fuzzy sets $\mathfrak{F} \subset [0, 1]^X$, the T-preorder $Inf_{\sigma \in \mathfrak{F}} J^{\alpha}_{T}$ has the elements in $\mathfrak{F}$ as T-states. If $\mu(x, y) = Inf_{\sigma \in \mathfrak{F}} J^{\sigma}_{T}(x, y)$, it follows $\mu \leq J^{\sigma}_{T}, \forall \sigma \in \mathfrak{F}$, or, equivalently $\sigma \in T(\mu), \forall \sigma \in \mathfrak{F}$, or $\mathfrak{F} \subset T(\mu)$. For example, with $X = [0, 1]$ and the two functions $\sigma_1(x) = x, \sigma_2(x) = 1 - x$, it is

$$\mu(x, y) = min(J^{\sigma_1}_{T}(x, y), J^{\sigma_2}_{T}(x, y))$$

a T-preorder. With $T = W$,

$$\begin{aligned} \mu(x, y) &= min(min(1, 1 - \sigma_1(x) + \sigma_1(y)), min(1, 1 - \sigma_2(x) + \sigma_2(y))) \\ &= min(1, 1 - x + y, 1 + x - y), \end{aligned}$$

is a W-preorder.

5.3 The Characterization of T-Indistinguishabilities

Let us consider, the T-indistinguishabilities

$$E^{\sigma}_{T}(x, y) = min(J^{\sigma}_{T}(x, y), J^{\sigma}_{T}(y, x)) = min(J_T(\sigma(x), \sigma(y)), J_T(\sigma(y), \sigma(x))).$$

If μ is a fuzzy relation in $[0, 1]^{X \times X}$, consider $T(\mu)$. Obviously

$$\mu(x, y) \leq \underset{\sigma \in T(\mu)}{Inf} \; E^{\sigma}_{T}(x, y), \text{ for all x, y in } X.$$

In the same vein that in the case of T-Preorders,

- *μ is a T-indistinguishability if and only if* $\mu = Inf_{\sigma \in T(\mu)} T_T^{\sigma}$.

  For example,

- $$T = min, \mu(x, y) = \underset{\sigma \in T(\mu)}{Inf} \; min\left(\begin{cases} 1, & \text{if } \sigma(x) \le \sigma(y) \\ \sigma(y), & \text{if } \sigma(x) > \sigma(y) \end{cases}, \begin{cases} 1, & \text{if } \sigma(y) \le \sigma(x) \\ \sigma(x), & \text{if } \sigma(y) > \sigma(x) \end{cases}\right)$$
$$= \underset{\sigma \in T(\mu)}{Inf} \begin{cases} 1, & \text{if } \sigma(x) = \sigma(y) \\ min(\sigma(x), \sigma(y)), & \text{otherwise} \end{cases}$$
$$= \begin{cases} 1, & \text{if } \sigma(x) = \sigma(y) \\ \underset{\sigma \in T(\mu)}{Inf} \; min(\sigma(x), \sigma(y)), & \text{otherwise} \end{cases}$$

- $$T = prod, \mu(x \cdot y) = \underset{\sigma \in T(\mu)}{Inf} \; min\left(\begin{cases} 1, & \text{if } \sigma(x) \le \sigma(y) \\ \frac{\sigma(y)}{\sigma(x)}, & \text{if } \sigma(x) > \sigma(y) \end{cases}, \begin{cases} 1, & \text{if } \sigma(y) \le \sigma(x) \\ \frac{\sigma(x)}{\sigma(y)}, & \text{if } \sigma(y) > \sigma(x) \end{cases}\right)$$
$$= \underset{\sigma \in T(\mu)}{Inf} \begin{cases} 1, & \text{if } \sigma(x) = \sigma(y) \\ min\left(\frac{\sigma(y)}{\sigma(x)}, \frac{\sigma(x)}{\sigma(y)}\right), & \text{otherwise} \end{cases}$$
$$= \begin{cases} 1, & \text{if } \sigma(x) = \sigma(y) \\ \underset{\sigma \in T(\mu)}{Inf} \; min\left(\frac{\sigma(y)}{\sigma(x)}, \frac{\sigma(x)}{\sigma(y)}\right), & \text{otherwise} \end{cases}$$

- $$T = W, \mu(x, y) = \underset{\sigma \in T(\mu)}{Inf} \; min(min(1, 1 - \sigma(x) + \sigma(y), 1 - \sigma(y) + \sigma(x)))$$
$$= \underset{\sigma \in T(\mu)}{Inf} \; min(1, 1 - max(\sigma(x) - \sigma(y), \sigma(y) - \sigma(x)))$$
$$= \underset{\sigma \in T(\mu)}{Inf} \; min(1, 1 - |\sigma(x) + \sigma(y)|)$$
$$= \underset{\sigma \in T(\mu)}{Inf} \; (1 - |\sigma(x) - \sigma(y)|)$$

Remark 5.3.1 Like in the case of T-Preorders, for any family of functions $\mathfrak{F} \subset [0, 1]^X$, the T-indistinguishability $Inf_{\sigma \in \mathfrak{F}} E_T^{\sigma}$ has the elements in $\mathfrak{F}$ as T-states. Notice that with $\mu(x, y) = Inf_{\sigma \in \mathfrak{F}} E_T^{\sigma}(x, y) = Inf_{\sigma \in \mathfrak{F}} min(J_T^{\sigma}(x, y), J_T^{\sigma}(y, x))$, it follows $\mu(x, y) \le J_T^{\sigma}(x, y)$, for all x, y in X, that is equivalent to $\sigma \in T(\mu)$. Hence,

$$\forall \sigma \in \mathfrak{F} \Rightarrow \sigma \in T(\mu), \text{ or } \mathfrak{F} \subset T(\mu).$$

With $X = [0, 1]$ and the two functions $\sigma_1(x) = x$, $\sigma_2(x) = 1 - x$, it is

$$\mu(x, y) = Min(E_T^{\sigma_1}(x, y), E_T^{\sigma_2}(x, y)),$$

a T-indistinguishability.

- With $T = W$, is

$$\mu(x, y) = Min(1 - |\sigma_1(x) - \sigma_1(y)|, 1 - |\sigma_2(x) - \sigma_2(y)|)$$

$$= Min(1 - |x - y|, 1 - |y - x|) = 1 - |x - y|.$$

Notice that with $\sigma_1(x) = x$, $\sigma_2(x) = x^2$, results

$$\mu(x, y) = Min(1 - |x - y|, 1 - |x^2 - y^2|),$$

that is, $\mu(x, y) = 1 - |x - y|$, provided $x + y \leq 1$, and $\mu(x, y) = 1 - |x^2 - y^2|$ if $x + y > 1$.

- With $T = prod$, results

$$\mu(x, y) = Min\left(\begin{cases} 1, & x \leq y \\ \frac{y}{x}, & x > y \end{cases}, \begin{cases} 1, & y \leq x \\ \frac{1-y}{1-x}, & y > x \end{cases}\right) = \begin{cases} 1, & x = y \\ \frac{y}{x}, & y < x \\ \frac{1-y}{1-x}, & y > x \end{cases}.$$

- With $T = min$, results

$$\mu(x, y) = min\left(\begin{cases} 1, & x \leq y \\ y, & x > y \end{cases}, \begin{cases} 1, & y \leq y \\ 1 - y, & y > x \end{cases}\right) = \begin{cases} 1, & x = y \\ y, & x \neq y \leq 1/2 \\ 1 - y, & x \neq y > 1/2 \end{cases}$$

Example 5.3.2 A finite example with $X = \{1, 2, 3, 4\}$. Take $\sigma_1(x) = \frac{x}{4}$, $\sigma_2(x) = 1 - \frac{x}{4}$, and $T = W$. It is

$$[J_W^{\sigma_1}] = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 3/4 & 1 & 1 & 1 \\ 1/2 & 3/4 & 1 & 1 \\ 1/4 & 1/2 & 3/4 & 1 \end{pmatrix}$$

Since, f.e.,

$$J_W^{\sigma_1}(1, 2) = min\left(1, 1 - \frac{1}{4} + \frac{2}{4}\right) = 1, J_W^{\sigma_1}(3, 1) = min\left(1, 1 - \frac{3}{4} + \frac{1}{4}\right) = 1/2,$$

$$J_W^{\sigma_1}(4, 1) = min\left(1, 1 - 1 + \frac{1}{4}\right) = 1/4, J_W^{\sigma_1}(3, 4) = min\left(1, 1 - \frac{3}{4} + 1\right) = 1, \text{etc.}$$

It is also

$$[J_W^{\sigma_2}] = \begin{pmatrix} 1 & 3/4 & 1/2 & 1/4 \\ 1 & 1 & 3/4 & 1/2 \\ 1/2 & 1 & 1 & 3/4 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

since, f.e.,

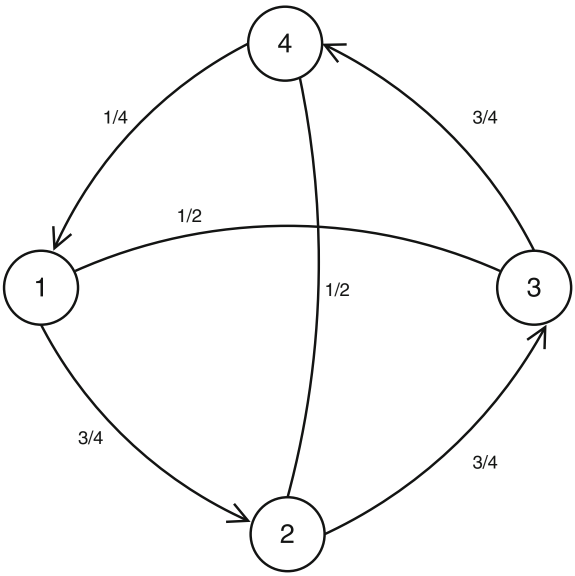
$$J_W^{\sigma_2}(1,2) = min\left(1, 1 - \frac{3}{4} + \frac{1}{2}\right) = 3/4, J_W^{\sigma_2}(3,1) = min\left(1, 1 - \frac{1}{4} + \frac{3}{4}\right) = 1/2,$$

$$J_W^{\sigma_2}(4,1) = min\left(1, 1 - 0 + \frac{3}{4}\right) = 1, J_W^{\sigma_2}(3,4) = min\left(1, 1 - \frac{1}{4} + 0\right) = 3/4, \text{etc.}$$

Hence,

$$[E_W] = min([J_W^{\sigma_1}], [J_W^{\sigma_2}]) = \begin{pmatrix} 1 & 3/4 & 1/2 & 1/4 \\ 3/4 & 1 & 3/4 & 1/2 \\ 1/2 & 3/4 & 1 & 3/4 \\ 1/4 & 1/2 & 3/4 & 1 \end{pmatrix}$$

is a W-indistinguishability on X, with the directed graph



Example 5.3.3 With an intersection $\mu \cdot \sigma = T \circ (\mu \times \sigma)$, ($T$ a continuous t-norm) define

$$E_T(\mu, \sigma) = Sup\,(\mu \cdot \sigma).$$

Obviously, $E_T(\mu, \sigma) = E_T(\sigma, \mu)$, since T is a commutative operation. To have,

$$E_T(\mu, \mu) = 1, \text{ or } Sup\, T \circ (\mu \times \mu) = 1,$$

since $T \circ (\mu \times \mu) \leq \mu$, it should be $Sup\, \mu = 1$, that is, since X is a finite set, μ should be a normalized fuzzy set, a fuzzy set for which it exists $x_0 \in X$ such that $\mu(x_0) = 1$.

From now on let's consider the set $\mathcal{N}(X) = \{\mu \in [0, 1]^X; Sup\, \mu = 1\}$. The mapping

$$E_T : \mathcal{N}(X) \times \mathcal{N}(X) \to [0, 1],$$

is reflexive and symmetric, that is, E_T is a similarity in $\mathcal{N}(X)$. Obviously, $E_T \leq E_{min}$, for all continuous t-norm T. Since,

$$\begin{aligned} T(E_T(\mu, \sigma), E_T(\sigma, \lambda)) &= T(Sup\, T(\mu, \sigma), Sup\, T(\sigma, \lambda)) \\ &= Sup\, T(T(\mu, \sigma), T(\sigma, \lambda)) \\ &\leq Sup\, T(\mu, \lambda) = E_T(\mu, \lambda), \end{aligned}$$

it results that, in addition, E_T is a T^*-indistinguishability for all continuous t-norm T^* such that $T^* \leq T$. In particular, E_{min} is not only a min-indistinguishability but a T-indistinguishability for all continuous t-norm T.

For example with $X = \{1, 2, 3, 4\}$, and the two fuzzy sets

$$\mu = 1|1 + 0.4|2 + 0.8|3 + 0.7|4, \sigma = 0.6|1 + 0.5|2 + 0.8|3 + 1|4$$

it follows

$$\mu \cdot \sigma = 0.6|1 + T(0.4, 0.5)|2 + T(0.8, 0.8)|3 + 0.7|4,$$

and

- $E_{min}(\mu, \sigma) = 0.8$, since $min(0.8, 0.8) = 0.8$, $min(0.4, 0.5) = 0.4$
- $E_{prod}(\mu, \sigma) = 0.64$, since $prod(0.8, 0.8) = 0.64$, $prod(0.4, 0.5) = 0.2$
- $E_W(\mu, \sigma) = 0.6$, since $W(0.8, 0.8) = 0.6$, $W(0.4, 0.5) = 0$

Remark 5.3.4 Provided $T = min$ or $T = prod_\varphi$, if $E_T(\mu, \sigma) > 0$, and $E_T(\sigma, \alpha) > 0$, it is

$$0 < T(E_T(\mu, \sigma), E_T(\sigma, \alpha)) \leq E_T(\mu, \alpha),$$

and $E_T(\mu, \alpha) > 0$. In these cases, E_T is said to be strictly transitive.

Nevertheless, in the case in which a fuzzy relation $E : X \times X \to [0, 1]$ is W_φ-transitive, from

$$E(x, y) > 0 \text{ and } E(y, z) > 0,$$

what follows is that, from the equivalence given by the existence of $r > 0$ such that $E(x, y) > r$, $E(y, z) > r$, it is

$$W_\varphi(r, r) = \varphi^{-1}(max(0, 2\varphi(r) - 1)) \leq W_\varphi(E(x, y), E(y, z)) \leq E(x, z).$$

Hence, to have $0 < E(x, z)$, it is necessary that $0 < W_\varphi(r, r)$, that is,

$$0 < max(0, 2\varphi(r) - 1), \text{ or } r > \varphi^{-1}(0.5).$$

For example, if it is $T = W$, it should be $r > 0.5$, and if $T = W_\varphi$ with $\varphi(x) = x^2$ it should be $r > \sqrt{0.5} = 0.7071$. In the case of the fuzzy relation

$$E(\mu, \sigma) = \frac{\sum_{i-1}^{n} min(\mu(x_i), \sigma(x_i))}{max(\sum_{i-1}^{n} \mu(x_i), \sum_{i-1}^{n} \sigma(x_i))},$$

with $X = \{x_1, \ldots, x_n\}$, it results W_φ-transitive for $\varphi(x) = x^2$. Thus, if

$$0.71 < E(\mu, \sigma), \text{ and } 0.71 < E(\sigma, \alpha)$$

it follows $0 < E(\mu, \alpha)$, since $W_\varphi(0.71, 0.71) = \sqrt{max(0, 2 \times 0.71^2 - 1)} = 0.09$.

Chapter 6
Fuzzy Arithmetic

6.1 Introduction

As it was explained before, any operation $* : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, can be extended to one $\circledast : [0, 1]^{\mathbb{R}} \times [0, 1]^{\mathbb{R}} \to [0, 1]^{\mathbb{R}}$, by means of the extension principle:

$$(\mu \circledast \sigma)(t) = \underset{t=x*y}{Sup}\ min(\mu(x), \sigma(y)).$$

This extension includes the crisp subsets $A \subset \mathbb{R}$, since $\mu_A \in \{0, 1\}^{\mathbb{R}} \subset [0, 1]^{\mathbb{R}}$. For example, with $A_1 = \{1, 2, 3\}$, and $A_2 = \{1, 3, 5\}$, and only taking into account the numbers in $\mathbb{N} \subset \mathbb{R}$, it is

$$(\mu_{A_1} \oplus \mu_{A_2})(t) = \underset{t=x+y}{Sup}\ min(\mu_{A_1}(x), \mu_{A_2}(y)), \text{ with } t, x, y \text{ in } \mathbb{N},$$

and $+$ the addition of natural numbers. Since $x + y \in \{2, 4, 5, 3, 5, 7, 8\}$, it results $\mu_{A_2} \oplus \mu_{A_2} = \mu_{\{2,3,4,5,6,7,8\}}$, that is $A_1 \oplus A_2 = \{2, 3, 4, 5, 6, 7, 8\}$.

In the same vein, if $A_1 = [a, b]$, and $A_2 = [c, d]$ are intervals of the real line $\mathbb{R}$, it results

$$\begin{aligned}(\mu_{A_1} \oplus \mu_{A_2})(t) &= \underset{t=x+y}{Sup}\ min(\mu_{[a,b]}(x), \mu_{[c,d]}(y)) \\ &= \mu_{[a+c,b+d]}, \text{ or } [a, b] \oplus [c, d] = [a + c, b + d].\end{aligned}$$

Analogously, it results

- $[a, b] \ominus [c, d] = [a - d, b - c]$
- $[a, b] \otimes [c, d] = [min(ad, ac, bd, bc), max(ad, ac, bd, bc)]$
- If $0 \notin [c, d]$, $[a, b] \oslash [c, d] = \left[min\left(\frac{a}{c}, \frac{a}{d}, \frac{b}{c}, \frac{b}{d}\right), max\left(\frac{a}{c}, \frac{a}{d}, \frac{b}{c}, \frac{b}{d}\right)\right]$.

E. Trillas and L. Eciolaza, *Fuzzy Logic*, Studies in Fuzziness and Soft Computing 320, DOI 10.1007/978-3-319-14203-6_6

For example,

- $[7, 8] \oplus [-1, 9] = [6, 17]$
- $[7, 8] \ominus [-1, 9] = [-2, 9]$
- $[3, 4] \odot [2, 2] = [6, 8]$
- $[4, 10] \oslash [1, 2] = [2, 10]$
- $2 \oplus [7, 8] = [2, 2] \oplus [7, 8] = [9, 10]$
- $2 \odot [7, 8] = [2, 2] \odot [7, 8] = [min(14, 16, 14, 16), max(14, 16)] = [14, 16]$
- $[7, 8] \oslash 2 = [7, 8] \oslash [2, 2] = \left[min\left(\frac{7}{2}, \frac{8}{2}\right), max\left(\frac{7}{2}, \frac{8}{2}\right)\right] = \left[\frac{7}{2}, 4\right]$

In short, and accordingly with the 'preservation of the classical case', through the extension principle both the numerical and the interval arithmetics are preserved. What follows is a yet larger arithmetic with fuzzy concepts.

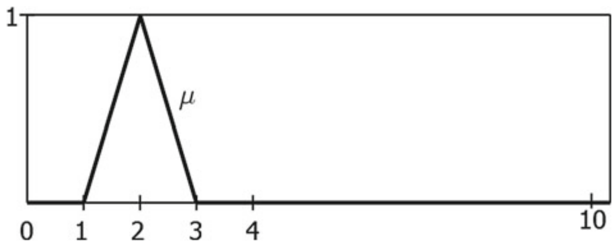
For all $\mu \in [0, 1]^{\mathbb{R}}$, and all $r \in [0, 1]$, it can be computed that:

- $$\begin{aligned}(r \oplus \mu)(t) = (\mu_r \oplus \mu)(t) &= \underset{t=x+y}{Sup}\ min(\mu_r(x), \mu(y)) \\ &= \underset{t=x+y}{Sup}\ min\begin{pmatrix}1, \mu(y), & \text{if } x = r \\ 0, \mu(y), & \text{if } x \neq r\end{pmatrix} = \underset{t=x+y}{Sup}\begin{pmatrix}\mu(y), & x = r \\ 0, & x \neq r\end{pmatrix} \\ &= \mu(t - r).\end{aligned}$$

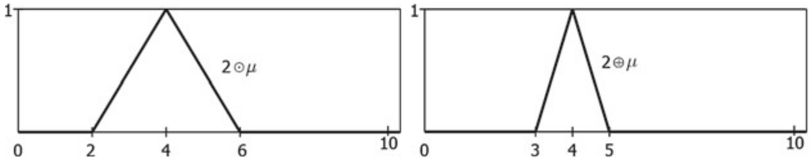
Hence $((-1) \oplus \mu)(t) = \mu(t - 1)$, $(\mu_1 \oplus \mu)(t) = \mu(t - 1)$, $(\mu_0 \oplus \mu)(t) = \mu(t)$.

- $(r \odot \mu)(t) = (\mu_r \odot \mu)(t) = \underset{t=x\cdot y}{Sup}\ min(\mu_r(x), \mu(t)) = \mu\left(\frac{t}{r}\right)$, if $r \neq 0$. Hence, $\left(\frac{1}{r} \odot \mu\right)(t) = \mu(rt)$, $(1 \odot \mu)(t) = \mu(t)$, $(\mu_1 \odot \mu)(t) = \mu(t)$.
- $(\mu_0 \odot \mu)(t) = \underset{t=x\cdot y}{Sup}\ min(\mu_0(x), \mu(t)) = 0 = \mu_0(t)$.
- $\frac{1}{\mu}(t) = \underset{t=\frac{x}{y}}{Sup}\, min(1, \mu(y)) = \underset{t=\frac{x}{y}}{Sup}\mu(y) = \begin{cases}\mu\left(\frac{1}{t}\right), & \text{if } t \neq 0 \\ 0, & \text{if } t = 0\end{cases}$

Example 6.1.1 Given the fuzzy set



compute $2 \odot \mu$ and $2 \oplus \mu$.
It is $(2 \odot \mu)(t) = \mu(t/2)$, and $(2 \oplus \mu)(t) = \mu(t - 2)$. Hence,



It should be pointed out that, since the cartesian product $\mu \times \sigma$ of $\mu \in [0, 1]^X$ and $\sigma \in [0, 1]^Y$, is defined by

$$\mu \times \sigma(x, y) = min(\mu(x), \sigma(y)), \mu \times \sigma \in [0, 1]^{X \times Y},$$

the operations $\circledast$, with $* \in \{+, -, \cdot, :\}$, can also be represented in the form

$$\mu \circledast \sigma(t) = \underset{t=x*y}{Sup}\, (\mu \times \sigma)(x, y),$$

that in the finite case are simply,

$$\mu \circledast \sigma(t) = \underset{t=x*y}{Max}(\mu \times \sigma)(x, y).$$

Example 6.1.2 If $X = \{1, 2, 3\}$, and $\mu = 0.7|1 + 0.9|2 + 1|3$, $\sigma = 0.8|1 + 0.9|3$, since $X \times X = \{(1, 1), (1, 2), \ldots, (3, 2), (3, 3)\}$, it results

$$\begin{aligned} \mu \times \sigma = {} & 0.7|(1, 1) + 0|(1, 2) + 0.7|(1, 3) + 0.8|(2, 1) \\ & + 0|(2, 2) + 0.9|(2, 3) + 0.8|(3, 1) + 0|(3, 2) + 0.9|(3, 3). \end{aligned}$$

Since $t = x + y$ will take the values $1 + 1 = 2$, $1 + 2 = 3$, $1 + 3 = 4$, $2 + 1 = 3$, $2 + 2 = 4$, $2 + 3 = 5$, $3 + 1 = 4$, $3 + 2 = 5$,$3 + 3 = 6$, that is $t \in \{2, 3, 4, 5, 6\}$, it results:

$$\mu \oplus \sigma(2) = 0.7, \mu \oplus \sigma(3) = max(0, 0.8) = 0.8, \mu \oplus \sigma(4) = max(0.7, 0.8, 0) = 0.8,$$

$$\mu \oplus \sigma(5) = max(0.9, 0) = 0.9, \mu \oplus \sigma(6) = 0.9$$

Hence,

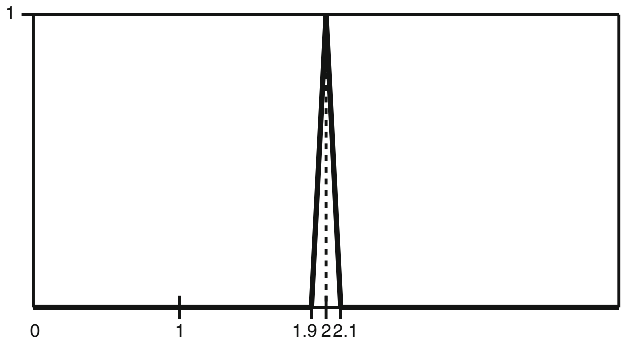
$$\mu \oplus \sigma = 0.7|2 + 0.8|3 + 0.8|4 + 0.9|5 + 0.9|6.$$

For the product $\mu \odot \sigma$, it is $t = x \cdot y \in \{1, 2, 3, 4, 6, 9\}$, and

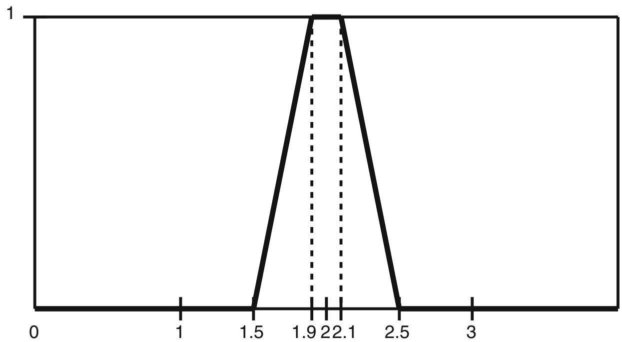
$$\mu \cdot \sigma = 0.7|1 + 0.8|2 + 0.8|3 + 0|4 + 0.9|6 + 0.9|9.$$

6.2 Fuzzy Numbers

In the usual scientific computation is rather unusual to consider exact numbers, as in the case of taking $\pi \approx 3.1416$, or $\sqrt{2} \approx 1.4142$. In some cases, an interval containing the number is considered, for example, $\sqrt{2} \in [1.412, 1.413]$. Fuzzy numbers are just a "fuzzification" of this last idea. For example, the fuzzy number "approximately 2" can be taken either as the fuzzy set,



or as the fuzzy set.

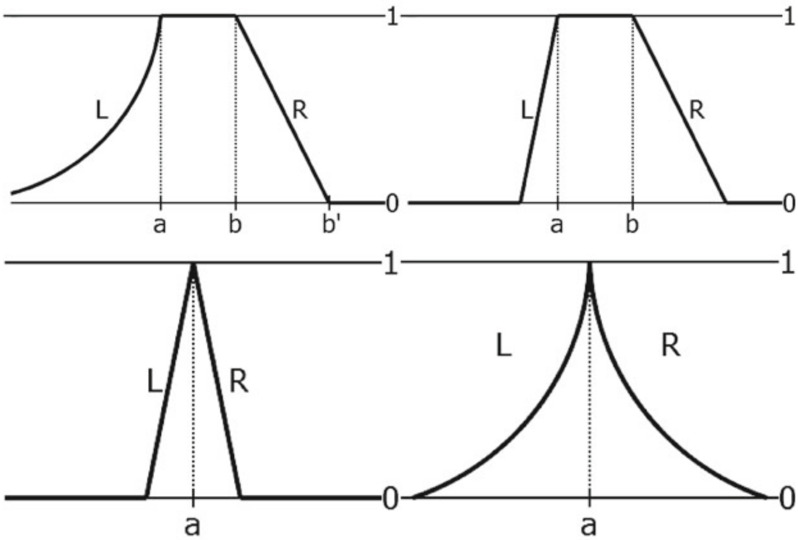


That kind of fuzzy sets are known as *fuzzy numbers*. Their definition is the following.

Definition 6.2.1 A fuzzy set $\mu \in [0, 1]^{\mathbb{R}}$ is a fuzzy number, provided there is a closed interval $[a, b] \in \mathbb{R}$, such that

1. $\mu(x) = 1$, for all $x \in [a, b]$
2. It exists a function $L : (-\infty, a) \to [0, 1]$, that is continuous and non-decreasing, such that $\mu(x) = L(x)$, for all $x \in (-\infty, a)$.
3. It exists a function $R : (b, +\infty) \to [0, 1]$, that is continuous and decreasing, such that $\mu(x) = R(x)$, for all $x \in (b, +\infty)$.

The following are examples of fuzzy numbers.



6.2.1 Operations with Fuzzy Numbers

It can be proven that, *if* $* \in \{+, -, \times, :\}$, *and if* μ, σ *are fuzzy numbers, then* $\mu \circledast \sigma$ *is also a fuzzy number*. For example, with the fuzzy number,

$$\mu_3(x) = \begin{cases} x - 2, & \text{if } \ \mathrm{x} \in (2, 3) \\ 4 - x, & \text{if } \ \mathrm{x} \in (3, 4) \\ 0, & \text{otherwise,} \end{cases}$$

for which $[a, b] = [3, 3] = \{3\}$, $L(x) = x - 2$, $R(x) = 4 - x$, it results

$$(\mu_3 \oplus \mu_3)(t) = \underset{t=x+y}{Sup}\ min(\mu_3(x), \mu_3(y)) = \begin{cases} \frac{t-4}{2}, & \text{if } \ t \in [4, 6] \\ \frac{8-t}{2}, & \text{if } \ t \in [6, 8] \\ 0, & \text{otherwise} \end{cases}$$

since for $t < 4$, or $x+y < 4$, and for $t > 8$, or $x+y > 8$, it should be $(\mu_3 \oplus \mu_3)(t) = 0$; for $t = 6$, or $x + y = 6$, it should be $(\mu_3 \oplus \mu_3)(6) = 1$, and:

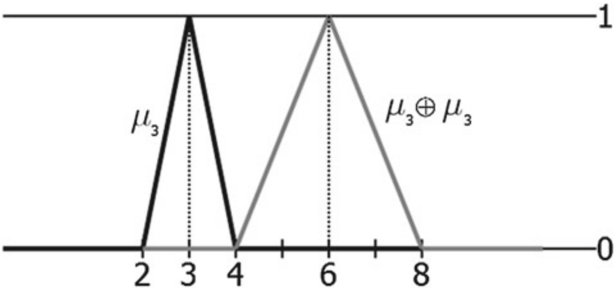
- For $t \in [4, 6]$, L is the segment joining (4, 0) and (6, 1), that is

$$0 = \begin{vmatrix} x & y & 1 \\ 4 & 0 & 1 \\ 6 & 1 & 1 \end{vmatrix} = -x + 2y + 4, \quad \text{or} \quad y = \frac{x-4}{2}$$

- For $t \in [6, 8]$, R is the segment joining (6, 1) and (8, 0), that is

$$0 = \begin{vmatrix} x & y & 1 \\ 6 & 1 & 1 \\ 8 & 0 & 1 \end{vmatrix} = x + 2y - 8, \quad \text{or} \quad y = \frac{8 - x}{2}$$

Graphically



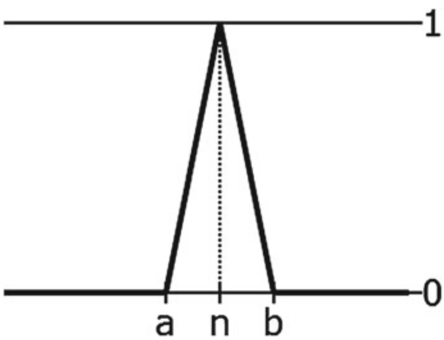
Notice that the result is also a fuzzy number μ_6, such that the amplitude of it ($[8, 4]$, $8 - 4 = 4$) is twice than that of the initial $\mu_3([2, 4], 4 - 2 = 2)$. When operating with fuzzy numbers, the amplitude grows. If there is uncertainty about the value 3, the uncertainty about $3 + 3$ is twice, and such uncertainty will grow each time we 'add' more numbers.

6.2.2 Operations with Triangular Fuzzy Numbers

Let's show a systematic way of operating with triangular fuzzy numbers, that is, those represented by

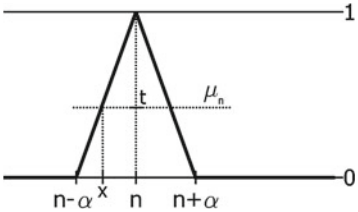
$$\mu_n(x) = \begin{cases} L(x), & \text{if } x \in (a, n) \\ R(x), & \text{if } x \in (n, b) \\ 0, & \text{otherwise,} \end{cases}$$

with L and R linear functions, and $\mu_n(n) = 1$. Graphically,



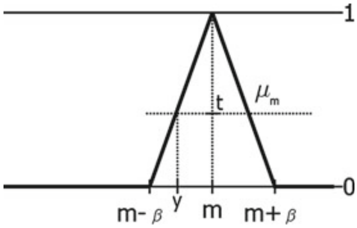
Namely, take

$$\mu_n(x) = \begin{cases} 1 + \frac{x-n}{\alpha}, & \text{if } x \in (n-\alpha, n) \\ 1 + \frac{n-x}{\alpha}, & \text{if } x \in (n, n+\alpha) \\ 0, & \text{otherwise,} \end{cases}$$



and

$$\mu_m(x) = \begin{cases} 1 + \frac{x-m}{\beta}, & \text{if } x \in (m-\beta, m) \\ 1 + \frac{m-x}{\beta}, & \text{if } x \in (m, m+\beta) \\ 0, & \text{otherwise,} \end{cases}$$



Obviously $(\mu_n \oplus \mu_m)(n+m) = 1$; if $t \leq m+n-(\alpha+\beta)$, $(\mu_n \oplus \mu_m)(t) = 0$; if $t \geq m+n+(\alpha+\beta)$, $(\mu_n \oplus \mu_m)(t) = 0$. Hence, the problem is reduced to compute the values of $\mu_n \oplus \mu_m$ in the two intervals $(m+n-(\alpha+\beta), m+n)$, and $(m+n, m+n+(\alpha+\beta))$. It can be done as follows:

1. If $x \in (n-\alpha, n)$, and $y \in (m-\beta, m)$, $t = x+y \in (n+m-(\alpha+\beta), n+m)$, and, since $\mu_n(x) = 1 + \frac{x-n}{\alpha}$, $\mu_m(y) = 1 + \frac{y+m}{\beta}$, it follows (with $\mu_n(x) = \mu_m(y) = z$): $x = \alpha z + n - \alpha$, $y = \beta z + m - \beta$, and

$$t = x + y = (\alpha+\beta)z + n + m - \alpha - \beta,$$

giving

$$z = \frac{t-(n+m)+(\alpha+\beta)}{\alpha+\beta},$$

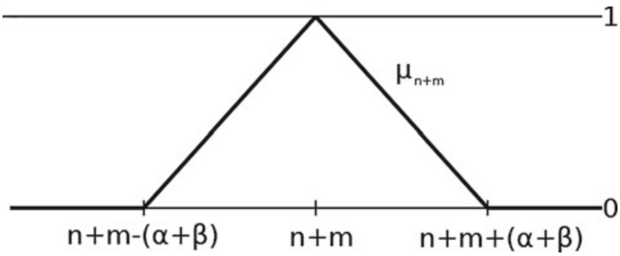
that is

$$(\mu_n \oplus \mu_m)(t) = \frac{t-(n+m)+(\alpha+\beta)}{\alpha+\beta}, \quad \text{if } \ t \in (m+n-(\alpha+\beta), m+n).$$

2. If $x \in (n, n+\alpha)$, and $y \in (m, m+\beta)$, $t \in (n+m, n+m+(\alpha+\beta))$, and like in (1), it is $z = \frac{t-(n+m)+(\alpha+\beta)}{\alpha+\beta}$, that is

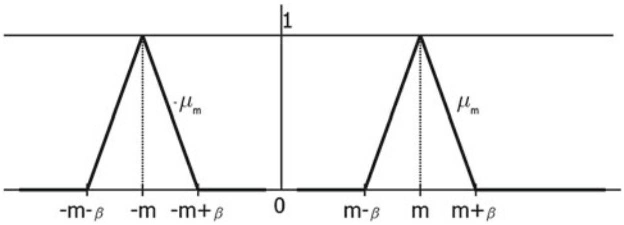
$$(\mu_n \oplus \mu_m)(t) = \frac{t-(n+m)-(\alpha+\beta)}{\alpha+\beta}, \quad \text{if } \ t \in (n+m, m+n-(\alpha+\beta)).$$

3. Graphically, with $\mu_n \oplus \mu_m = \mu_{n+m}$



For example, μ_3 with $\alpha = 1$, and μ_8 with $\beta = 2$, give $\mu_3 \oplus \mu_8 = \mu_{11}$ with $\alpha + \beta = 3$, that is the triangle on $[11 - 3, 11 + 3] = [8, 14]$.

Let's take into account the substraction $\mu_n \ominus \mu_m = \mu_n \ominus [(-1) \odot \mu_m]$, with $((-1) \odot \mu_m)(t) = \mu_m \left(\frac{t}{-1}\right) = \mu_m(-t)$, that means



with $-\mu_m = (-1) \odot \mu_m$. The fuzzy number $-\mu_m$ is obtained by translating μ_m symmetrically at the other side of $x = 0$. For example, with

$$\mu_3(x) = \begin{cases} x - 2, & \text{if } x \in (2, 3) \\ 4 - x, & \text{if } x \in (3, 4) \\ 0, & \text{otherwise,} \end{cases}$$

it results

$$(\mu_3 \ominus \mu_3)(t) = \left\{ \begin{array}{ll} \frac{t+2}{2}, & \text{if } x \in (-2, 0) \\ \frac{2-t}{2}, & \text{if } x \in (0, 2) \\ 0, & \text{otherwise.} \end{array} \right\}$$
$$= \mu_0(t), \text{ with the interval } [a, b] = [-2, +2].$$

For what concerns the product $\mu_n \odot \mu_m$ of two triangular fuzzy numbers, the method is the same of the sum, but remembering that since $(\mu_n \odot \mu_m)(t) = \underset{t=x\cdot y}{Sup\ min}(\mu_n(s), \mu_m(y))$, the value t will be reached by the product $x \cdot y$. For example, with μ_3 as below, let's compute $\mu_3 \odot \mu_3$. The process is as follows.

1. If $t \leq 4$, and $x \cdot y = t$, either $x \leq 2$, or $y \leq 2$. Hence, $(\mu_3 \odot \mu_3)(t) = 0$. Analogously, if $t \geq 16$, $(\mu_3 \odot \mu_3)(t) = 0$.
   Obviously, $(\mu_3 \odot \mu_3)(9) = 1$.
2. Hence, L will be defined in [4, 9], and R in [9, 16].
3. If $x, y \in [2, 3]$, it is $x \cdot y \in [4, 9]$, hence,

$$L(t) = \underset{t=x\cdot y}{Sup}\, min(x-2, y-2) = \alpha \Rightarrow \alpha = x-2,\ \alpha = y-2$$
$$\Rightarrow x = \alpha + 2,\ y = \alpha + 2 \Rightarrow t = x \cdot y = (\alpha+2)^2 \Rightarrow \alpha = \sqrt{t} - 2.$$

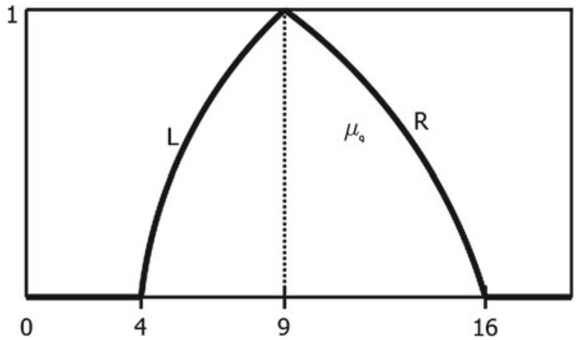
4. If $x, y \in [3, 4]$, it is $x \cdot y \in [9, 16]$, hence,

$$R(t) = \underset{t=x\cdot y}{Sup}\, min(4-x, 4-y) = \alpha \Rightarrow \alpha = 4-x,\ \alpha = 4-y$$
$$\Rightarrow x = 4-\alpha,\ y = 4-\alpha \Rightarrow t = (4-\alpha)^2 \Rightarrow \alpha = 4 - \sqrt{t}.$$

Finally,

$$(\mu_3 \odot \mu_3)(t) = \left\{\begin{array}{ll} \sqrt{t}-2, & \text{if } t \in [4, 9] \\ 4-\sqrt{t}, & \text{if } t \in [9, 16] \\ 0, & \text{otherwise.} \end{array}\right\} = \mu_9(t).$$

Graphically



Remark 6.2.2 It should be pointed out that, in the case of the product of triangular fuzzy numbers, the result is not a triangular (linear) fuzzy number, since functions L and R are not linear. In addition the interval $[a, b]$ is not symmetrical. Look that for μ_9 is [4, 16] whose mid point is not 9 but 10.

Let's now consider the quotient $\oplus$ of fuzzy numbers defined by

$$(\mu_n \oplus \mu_m)(t) = \underset{t=\frac{x}{y}}{Sup}\, min(\mu_n(x), \mu_m(y)),$$

or $\mu_n \oplus\!\!\!\!\div \mu_m = \mu_n \odot \frac{1}{\mu_m}$, with $\left(\frac{1}{\mu_m}\right)(x) = \begin{cases} \mu_m\left(\frac{1}{x}\right), & \text{if } x \neq 0 \\ 0, & \text{if } x = 0 \end{cases}$, and $\left(\frac{1}{\mu_m}\right)\left(\frac{1}{m}\right) = 1$.

Let's compute the example $\mu_3 \div \mu_3$, with μ_3 as in Sect. 6.2.2. It will be:

$$\left(\frac{1}{\mu_3}\right)(y) = \begin{cases} 4 - \frac{1}{y}, & \text{if } y \in \left(\frac{1}{4}, \frac{1}{3}\right) \\ \frac{1}{y} - 2, & \text{if } y \in \left(\frac{1}{3}, \frac{1}{2}\right) \\ 0, & \text{otherwise.} \end{cases}$$

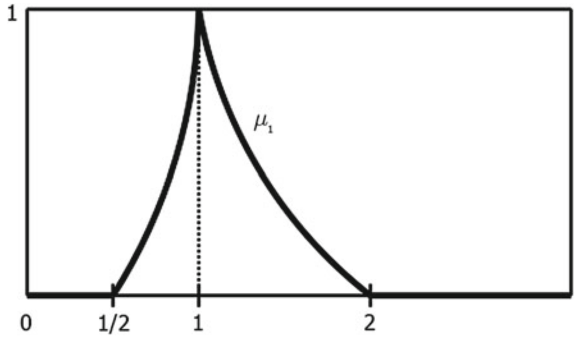
Hence, for $\mu_3 \div \mu_3 = \mu_3 \odot \frac{1}{\mu_3}$, it will be $(\mu_3 \div \mu_3)(1) = 1$, and

- L, $\alpha = 4 - \frac{1}{y} = x - 2 \Rightarrow t = x \cdot y = \frac{\alpha+2}{4-\alpha} \Rightarrow \alpha = \frac{4t-2}{t+1}$
- R, $\alpha = 4 - x = \frac{1}{y} - 2 \Rightarrow t = x \cdot y = \frac{4-\alpha}{\alpha+2} \Rightarrow \alpha = \frac{4-2t}{t+1}$.

That is

$$(\mu_3 \div \mu_3)(t) = \left\{ \begin{array}{ll} \frac{4t-2}{t+1}, & \text{if } t \in [\frac{1}{2}, 1] \\ \frac{4-2t}{t+1}, & \text{if } t \in [1, 2] \\ 0, & \text{otherwise} \end{array} \right\} = \mu_1(t)$$

Graphically,



Remark 6.2.3 Obviously, like in the product's case, the result is not a linear-triangular fuzzy number, and the basic interval $[a, b]$ is not symmetrical respect to the point $\frac{n}{m}$.

6.2.3 Note

As it is easy to see, if $\mu, \sigma \in [0, 1]^{\mathbb{R}}$, and $* \in \{+, -, \cdot, :\}$, it follows the relation between the corresponding α-cuts:

$$(\mu \circledast \sigma)_\alpha = \mu_\alpha \circledast \sigma_\alpha,$$

with the particular supposition that when $* = :$, it should be required that $0 \notin \mu_\alpha$ for all $\alpha \in (0, 1]$. Notice that it also follows

$$\mu \circledast \sigma = \bigcup_{\alpha \in [0,1]} \alpha \wedge (\mu * \sigma)_\alpha = \bigcup_{\alpha \in [0,1]} \alpha \wedge [\mu_\alpha \circledast \sigma_\alpha].$$

It is for this equality that the before hand computations were made. For example, with

$$(\mu_1)(x) = \begin{cases} \frac{x+1}{2}, & \text{if } t \in (-1, 1) \\ \frac{3-x}{2}, & \text{if } t \in (1, 3) \\ 0, & \text{otherwise} \end{cases}, \quad (\mu_3)(x) = \begin{cases} \frac{x-1}{2}, & \text{if } t \in (1, 3] \\ \frac{5-x}{2}, & \text{if } t \in (3, 5) \\ 0, & \text{otherwise} \end{cases}$$

the corresponding α-cuts are

$$(\mu_1)_\alpha = [2\alpha - 1, 3 - 2\alpha]\ ,\ (\mu_3)_\alpha = [2\alpha + 1, 5 - 2\alpha],$$

with which

$$[\mu_1 \oplus \mu_3]_\alpha = [4\alpha, 8 - 4\alpha], \text{for}\, \alpha \in (0, 1].$$

Hence,

$$(\mu_1 \oplus \mu_3)(t) = \begin{cases} \frac{x}{4}, & \text{if } t \in (0, 4] \\ \frac{8-x}{4}, & \text{if } t \in (4, 8] = \mu_4(t) \\ 0, & \text{otherwise} \end{cases}$$

6.3 A Note on the Lattice of Fuzzy Numbers

As it is well known, $(\mathbb{R}, min, max)$ is a distributive lattice that come from the totally ordered set $(\mathbb{R}, \leq)$. The order $\leq$ is definable from the lattice operations *min*, *max* by

$$a \leq b \Leftrightarrow a = min(a, b) \Leftrightarrow b = max(a, b).$$

In addition, with $a' = 1 - a$, it is $min(a, b) = (max(a', b'))'$, and $max(a, b) = (min(a', b'))'$. Let's extend these operations to the set $\mathbb{R}^*$ of all fuzzy numbers, by

$$(\mu \owedge \sigma)(t) = \underset{t=min(x,y)}{Sup} (\mu \times \sigma)(x, y), \text{ and } (\mu \ovee \sigma)(t) = \underset{t=max(x,y)}{Sup} (\mu \times \sigma)(x, y).$$

Without the proof, let's state:

Theorem 6.3.1 $(\mathbb{R}^*, \owedge, \ovee)$ *is a distributive lattice, that means:*

- $\mu \owedge \sigma,\ \mu \ovee \sigma \in \mathbb{R}^*$
- $\mu \owedge \sigma = \sigma \owedge \mu,\ \mu \owedge \mu = \mu$
- $\mu \ovee \sigma = \sigma \ovee \mu,\ \mu \ovee \mu = \mu$
- $\mu \owedge (\mu \ovee \sigma) = \mu,\ \mu \ovee (\mu \owedge \sigma) = \mu$
- $\mu \owedge (\sigma \ovee \lambda) = (\mu \owedge \sigma) \ovee (\mu \owedge \lambda)$
- $\mu \ovee (\sigma \owedge \lambda) = (\mu \ovee \sigma) \owedge (\mu \ovee \lambda),$

for all μ, σ, λ *in* $\mathbb{R}^*$.

Hence, $\mathbb{R}^*$ *can be endowed with the* partial *order given by*

$$\mu \leq^* \sigma \Leftrightarrow \mu \owedge \sigma = \mu \Leftrightarrow \mu \ovee \sigma = \sigma.$$

6.3.1 Example

With $X = \{1, 2, 3\}$, and $\mu = 0.8|1 + 0.7|2 + 1|3,\ \sigma = 0.9|1 + 1|2 + 0.6|3$, compute $\mu \owedge \sigma, \mu \ovee \sigma$.

Since, $t = min(x, y)$ and $t = max(x, y)$ belong to $\{1, 2, 3\}$, it results:

- $(\mu \owedge \sigma)(t) = \underset{t=min(x,y)}{Max} [min(\mu(x), \sigma(y))]$, and:
  - $(\mu \owedge \sigma)(1) = max(min(\mu(1), \sigma(1)), min(\mu(1), \sigma(2)), min(\mu(2), \sigma(1)),$ $min(\mu(1), \sigma(3)), min(\mu(3), \sigma(1)) = max(min(0.8, 0.9), min(0.8, 1),$ $min(0.7, 0.9), min(0.8, 0.6), min(1, 0.9)) = max(0.8, 0.8, 0.7, 0.6, 0.9) = 0.9$
  - $(\mu \owedge \sigma)(2) = max(min(\mu(2), \sigma(3)), min(\mu(3), \sigma(2)))$ $= max(min(0.7, 0.6), min(1, 1)) = max(0.6, 1) = 1$
  - $(\mu \owedge \sigma)(3) = min(\mu(3), \sigma(3)) = min(1, 0.6) = 0.6$

  that is $\mu \owedge \sigma = 0.9|1 + 1|2 + 0.6|3$.
  This fuzzy set is different from $\mu \cdot \sigma = 0.8|1 + 0.7|2 + 0.6|3$, with the t-norm *min*.
- $(\mu \ovee \sigma)(t) = \underset{t=max(x,y)}{Max} [min(\mu(x), \sigma(y))]$, and:
  - $(\mu \ovee \sigma)(1) = min(\mu(1), \sigma(1)) = min(0.8, 0.9) = 0.8$
  - $(\mu \ovee \sigma)(2) = max(min(\mu(1), \sigma(2)), min(\mu(2), \sigma(1)))$ $= max(min(0.8, 1), min(0.7, 0.9)) = max(0.8, 0.7) = 0.8$

– $(\mu \ovee \sigma)(3) = max(min(\mu(1), \sigma(3)), min(\mu(3), \sigma(1)), min(\mu(2), \sigma(3)),$ $min(\mu(3), \sigma(2)), min(\mu(3), \sigma(3))) = max(min(0.8, 0.6), min(1, 0.9),$ $min(0.7, 0.6), min(1, 1), min(1, 0.6)) = 1,$

that is $\mu \ovee \sigma = 0.8|1 + 0.8|2 + 1|3$.
This fuzzy set is different from $\mu + \sigma = 0.9|1 + 1|2 + 1|3$, with $+$ the t-conorm max.

Remark 6.3.2 It is easy to check that, although *it is not* $\mu \leq \sigma$ pointwise, *it is* $\mu \owedge \sigma \leq^* \mu \ovee \sigma$.

Remark 6.3.3 Since $t = min(x, y)$ means

- $t = x$, if $x \leq y$,
- $t = y$, if $y \leq x$,

it is immediate that

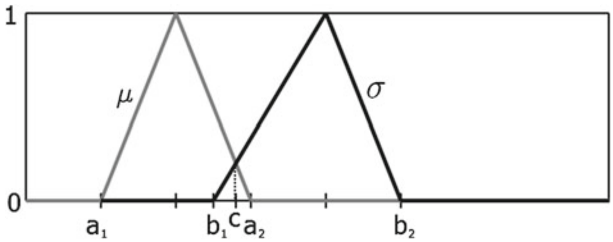
$$(\mu \owedge \sigma)(t) = max[\underset{t \leq x}{Sup}\, min(\mu(x), \sigma(t)), \underset{t \leq y}{Sup}\, min(\mu(t), \sigma(y))],$$

a formula that facilitates to obtain $\mu \owedge \sigma$, given μ and σ. Analogously, since $t = max(x, y)$ means

- $t = x$, if $y \leq x$,
- $t = y$, if $y \geq x$,

there is a similar formula for $\mu \ovee \sigma$.

For example, in the case of μ and σ given in the figure



it results:

1. If $t \leq a_1$, $(\mu \owedge \sigma)(t) = 0$
2. If $a_2 \leq t$, $(\mu \owedge \sigma)(t) = 0$
3. If $a_1 \leq t \leq b_1$, $(\mu \owedge \sigma)(t) = \mu(t)$
4. If $b_1 \leq t \leq c$, $(\mu \owedge \sigma)(t) = \sigma(t)$
5. If $c \leq t \leq a_2$, $(\mu \owedge \sigma)(t) = \mu(t)$.

Hence, $\mu \owedge \sigma = \mu$, and $\mu \leq^* \sigma$, although it is not $\mu \leq \sigma$.

6.4 A Note on Fuzzy Quantifiers

In classical logic, only two quantifiers are considered. The universal quantifier $\forall$ (for all), and $\exists$ (it exists, or for some), the existential quantifier, with the addition of $\exists!$ (exists only one). For example, given a sequence of real numbers (a_n), it is said that the real number a is its *limit*, when

$$\forall \epsilon > 0, \exists k \in \mathbb{N}, \forall n \in \mathbb{N} : [(n > k) \rightarrow (|a_n - a| < \epsilon)].$$

Analogously, a function $f : [a, b] \rightarrow \mathbb{R}$ is *bounded*, when

$$\exists M \in \mathbb{R}^+, \forall x \in [a, b] : [|f(x)| < M].$$

The importance of these two quantifiers to clearly write mathematical expressions does not need to be stressed. Nevertheless, both in arithmetic computing and in natural language more quantifiers are needed and used.

Examples of arithmetical quantifiers are the percentages. For example

- The 85 % of the employees are married
- Between the 40 and the 70 % of the employees are single.

For example, if it is known that

$$\begin{array}{l} \text{The 35\% of the employees are married} \\ \text{The 25\% of the married employees are young} \\ \hline \text{What can be said on the employees that are young?} \end{array}$$

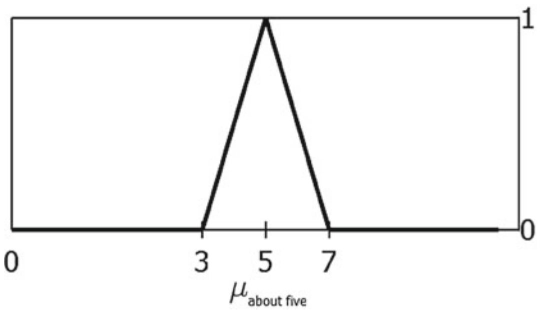
A question which answer is, obviously, $35 \times 25 = 875$, that is, at least the 8.75 % of the employees are young.

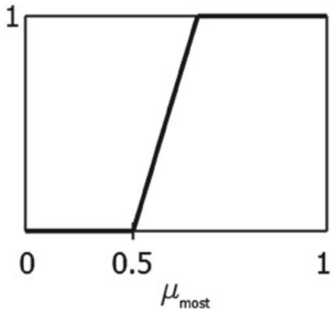
Another example is

$$\begin{array}{l} \text{Between 15 and 25 employees are married} \\ \text{Between 5 and 10 married employees are young} \\ \hline \text{What can be said on the employees that are young?} \end{array}$$

in which what matters is the length of the two intervals $[15, 25]$ $(l = 10)$, and $[5, 10]$ $(l = 5)$, that give at least, the interval $[5, 10 + (10 - 5)] = [5, 15]$ of employees that are young.

In natural language imprecise quantifiers like 'about five', 'about half', 'most', etc., appear and can be represented by means of fuzzy numbers. For example,





These fuzzy quantifiers are of two main types

- *Absolute quantifiers*, when are fuzzy numbers in $\mathbb{R}$ (independent of the cardinality of the universe of discourse)
- *Relative quantifiers*, when are fuzzy numbers in $[0, 1]$ (dependent of the cardinality of the universe of discourse)

Of course, the "interval" arithmetical quantifiers belong to the class of absolute (fuzzy) quantifiers as a (crisp) particular case, and the "percentage" arithmetical quantifiers belong to the class of relative (fuzzy) quantifiers as a (crisp) particular case.

The important problem is to compute the degree of validity, or truth, of statements with fuzzy quantifiers, like for example:

- In a given class, there are about three students whose fluency in English is low,

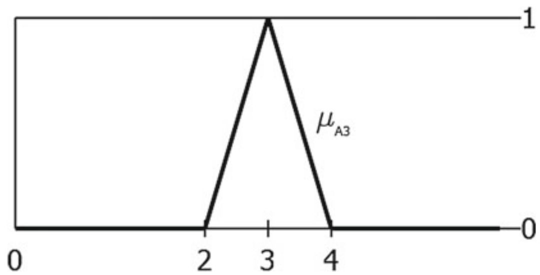
that firstly should be represented in fuzzy terms. To do that, call μ_{A_3} the fuzzy quantifiers, and μ_{LF} the degree of low English-fluency, and rewrite the above statement as

- There are μ_{A_3}, μ_{LF}'s.

Then, the degree of validity can be taken as

$$t = \mu_{A_3}(|\mu_{LF}|), \text{ with } |\mu_{LF}| \text{ the cardinality of } \mu_F.$$

For example, with



and the scores of LF of the (supposedly) five students $\{1, 2, 3, 4, 5\}$ in the class, given by

$$\mu_{LF} = 0|1 + 0|2 + 0.75|3 + 1|4 + 0.5|5,$$

it is $|\mu_L| = 1 + 0.75 + 0.5 = 2.25$, and

$$t = \mu_{A_3}(2.25) = 2.25 - 2 = 0.25,$$

since the equation of the line joining the points (2, 0) and (3, 1), in the figure of μ_{A_3}, is $y = x - 2$.

6.4.1 Quantified Fuzzy Statements

Another, more general, kind of quantified fuzzy statements, is

$$\text{There are } Q \in X\text{, such that “}F_1(x)\text{is } P_1\text{”}, \ldots, \text{“}F_n(x) \text{ is } P_n\text{”},$$

with X the universe of discourse, $F_i : X \to F_i(X) \subset \mathbb{R}$, $1 \leq i \leq n$, and P_i a predicate in $F_i(X)$, $1 \leq i \leq n$. For example,

There are about 6 employees in the company that are young and whose computer skills are high.

where $X = \{x_1, \ldots, x_n\}$ is the set of employees, $Q =$ about 6, $F_1 =$ age, and $F_2 =$ computer skills. These statement can be compressed to the form

$$\text{There are } QH_1\text{'s } H_2\text{'s } \equiv \text{There are } Q(H_1 \text{ and } H_2),$$

with $H_1(x) = \mu_{P_1}(H_1(x))$, $H_2(x) = \mu_{P_2}(H_2(x))$, for all $x \in X$, and that correspond with the rewritting:

There are about 6 employees that are young and with high computer skills,

of the given statement. Finally, with $Z = |H_1 \cap H_2| = \sum_{i=1}^{n} min(\mu_{P_1}(F_1(x_i)), \mu_{P_2}(F_2(x_i)))$, and $Q(Z) = Q(|H_1 \cap H_2|)$, results the more compressed form

$$Z \text{ is } Q$$

corresponding to:

The number of employees that are young and with high computer skills, is about 6,

and that gives the truth-value $t = Q(Z)$

The compression of fuzzy quantified statements is essential for its representation and for computing its truth-value.

Let's consider the case with relative quantifiers. The most simple case is

- Among the $x \in X$, Q are such that "$F(x)$ is P",

for example,

- Among the company's employees almost all are young.

This statement can be compressed to

- Q are H's, with $H(x) = \mu_P(F(x))$,

corresponding to

- Almost all company's employees are young.

Finally, the last compression can be done in the form:

$$Z \text{ is } Q,$$

with $Z = \frac{|H|}{|X|}$, that gives the truth value $t = Q(Z)$.

Example 6.4.1 (A problem of inference) Given n statements of the compressed form "Z_i is Q_i", $1 \leq i \leq n$ with Q_i absolute or relative quantifiers, which statement of the form "Z_i is Q" can be inferred?

For example,

$$\begin{array}{rl} & \text{There are about 10 workers in the establishment} \\ & \text{About half of the establishment workers are women} \\ \hline \text{We search Q such that:} & \text{There are Q women in the stablishment} \end{array}$$

Now, it is Q_1 = about 10, Q_2 = about half, X_1 = set of workers, X_2 = set of women $\subset X_1$. Then the syllogism can be stated by

$$\begin{array}{l} Z_1 \text{ is } Q_1 \\ Z_2 \text{ is } Q_2 \end{array}$$

where $Z_1 = |X_1|$, $Z_2 = \frac{|X_1 \cup X_2|}{|X_1|} = \frac{|X_2|}{|X_1|}$. Hence the conclusion is "Z is Q", with $Z = |X_2|$, and it rests to compute Q.

A rule often used to obtain Q is the following. If its exists $f : \mathbb{R}^n \to \mathbb{R}$ such that $Z = f(Z_1, \ldots, Z_n)$, take $Q = f(Q_1, \ldots, Q_n)$. In the example, it is $Z_1 \cdot Z_2 = |X_2| = Z$, hence $Q = Q_1 \cdot Q_2$, and the inference is

$$Z_1 \cdot Z_2 \text{ is } Q_1 \cdot Q_2, \textit{with } \mu_{Q_1} \cdot \mu_{Q_2} = \textit{min}(\mu_{Q_1}, \mu_{Q_2}).$$

As a final example,

$$\begin{array}{l} \text{Most of the workers are young} \\ \text{About half of the young workers are women} \\ \hline \text{Most} \times \text{ About half workers are young women.} \end{array}$$

Chapter 7
Fuzzy Measures

7.1 Introduction

Fuzzy sets not only appear by representing imprecise predicates, but also partial or incomplete information. This is the case, for example, of a function X taking values in [0, 10], of which it is only known that

$$5 \leqslant X \leqslant 7, 8 \nleqslant X, \quad \text{and} \quad X \nleqslant 2.$$

This information can be represented by means of a fuzzy set $\mu_X \in [0, 10]^{\Omega}$, by taking:

- $\mu_X(x) = 1 \Leftrightarrow X$ takes the value x
- $\mu_X(x) = 0 \Leftrightarrow X$ don't takes the value x
- $\mu_X(x) = a \cdot x + b \Leftrightarrow$ It is unknown if X can take the value x,

and the above information resumed by:

$$\mu_X(x) = \begin{cases} 0, & \text{if} \quad x \in [0, 2] \cup [8, 10] \\ \frac{x-2}{3}, & \text{if} \quad x \in [2, 5] \\ 1, & \text{if} \quad x \in [5, 7] \\ 8 - x, & \text{if} \quad x \in [7, 8] \end{cases}$$

In these cases, the questions of the uncertainty concerning the questions,

- $X \in A$ for $A \subset [0, 10]$
- $X \underset{r}{\in} \underset{\backsim}{A}$, for $\mu_A \in [0, 10]^X$,

arise and suggests the problem of its measuring. In this section, several "measures" of that kind of uncertainty, including probabilities when applicable, are introduced and studied.

The area of a polygon is a measure of its extensional size, and the length of a segment is a measure of its longitudinal size. Everybody knows several examples of

E. Trillas and L. Eciolaza, *Fuzzy Logic*, Studies in Fuzziness
and Soft Computing 320, DOI 10.1007/978-3-319-14203-6_7

measures, as it is, for example, the number of apples in a basket. But, how can the concept of measure be formalized?

7.2 The Concept of a Measure

Given a set X, and provided

- $\mathfrak{F} \subset [0, 1]^X$ is a family of fuzzy subsets of X, such that $\mu_0,\ \mu_1 \in \mathfrak{F}$, and that $\preceq$ is a preorder in $\mathfrak{F}$ translating a qualitative binary relation between the elements in $\mathfrak{F}$,
- $(L, \leqslant)$ is a preordered set with first element 0,

we will say that $m : \mathfrak{F} \to L$ is a $(L, \leqslant)$-measure, whenever

1. $m(\mu_0) = 0$
2. If $\mu \preceq \sigma$, then $m(\mu) \leqslant m(\sigma)$.

Example 7.2.1 Take $\mathfrak{F} = [0, 1]^X$, $(L, \leqslant) = ([0, 1], \leqslant)$ the unit interval with the partial linear order $\leqslant$ of the real line, and the qualitative relation "μ *is less fuzzy than* σ", translated by the so-called sharpened order $\leqslant_S$ $(=\preceq)$ defined by

$$\mu \leqslant_S \sigma \Leftrightarrow \begin{cases} \mu(x) \leqslant \sigma(y), & \text{if} \quad \sigma(x) \leqslant 1/2 \\ \sigma(x) \leqslant \mu(x), & \text{if} \quad \sigma(x) > 1/2, \end{cases}$$

that is a reflexive, transitive and antisymmetric, crisp relation. The fuzzy set $\mu_{0.5}$ is the highest one, and all crisp sets $\mu \in \{0, 1\}^X$ are minimals in $(\mathfrak{F}, \leqslant_S)$. Any mapping, $m : [0, 1]^X \to [0, 1]$ such that

- If μ is crisp, then $m(\mu) = 0$
- $m(\mu_{0.5}) = 1$
- If $\mu \leqslant_S \sigma$, then $m(\mu) \leqslant m(\sigma)$,

is a $([0, 1], \leqslant)$-measure since $m(\mu_0) = 0$ because of $\mu_0 \in \{0, 1\}^X$. These measures are called *measures of fuzziness*, or *fuzzy entropies*.

If $X = \{x_1, \ldots, x_n\}$ is finite, the following mappings are examples of fuzzy entropies:

1. $m(\mu) = 1 - 2 \max\limits_{1 \leqslant i \leqslant n} |\mu(x_i) - \frac{1}{2}|$
2. $m(\mu) = 2 \max\limits_{1 \leqslant i \leqslant n} \mu(x_i) \cdot (1 - \mu(x_i))$
3. $m(\mu) = \sum_{i=1}^{n} \sigma(\mu(x_i))$, with $\sigma(x) = x \ln x - (1 - x) \ln(1 - x)$ (logarithmic entropy).
4. $m(\mu) = \frac{1}{2n} \sum_{1=1}^{n} |\mu(x_i) - \mu_{C_\mu}(x_i)|$, with $\mu_{C_\mu}(x) = \begin{cases} 1, \text{if } \mu(x) > 0.5 \\ 0, \text{if } \mu(x) \leqslant 0.5 \end{cases}$, the *closest crisp set to* μ (linear index of fuzziness)
5. $m(\mu) = \frac{1}{2n} \sqrt{\sum_{i=1}^{n} (\mu(x_i) - \mu_{C_\mu}(x_i))^2}$ (quadratic index of fuzziness).

Remark 7.2.2 1. For some specific problems, measures of fuzziness are selected verifying the additional property of symmetry:

- For some negation N, it is $m(\mu) = m(N \circ \mu) = m(\mu')$.

For example, with $N = 1 - id$, measures 1, 2, 3, 4 do verify this property of symmetry $m(\mu) = m(1 - \mu)$.

2. With each measure of fuzziness m, it can be defined a *measure of booleanity* $1 - m$, that, obviously verifies:

- the value of $1 - m$ is 1 for the crip sets.
- the value of $1 - m$ is 0 for $\mu_{1/2}$.
- $1 - m$ is non-increasing with respect to the order $\leqslant_S$.

Example 7.2.3 Take $\mathfrak{F} \in [0, 1]^X$, with the partial pointwise order '$\mu \leqslant \sigma \Leftrightarrow \mu(x) \leqslant \sigma(x)$, for all $x \in X$' (and such that $\mu_0,\ \mu_1 \in \mathfrak{F}$). A mapping $m : [0, 1]^X \to [0, 1]$ is a *fuzzy measure* provided m verifies:

1. $m(\mu_0) = 0$
2. $m(\mu_1) = 1$
3. If $\mu \leqslant \sigma$, then $m(\mu) \leqslant m(\sigma)$.

When $\mathfrak{F} \in \{0, 1\}^X \approx \mathbb{P}(X)$, a fuzzy measure is defined by

1. $m(\varnothing) = 0$
2. $m(X) = 1$
3. If $A \subset B$, then $m(A) \leqslant m(B)$.

For example, if X is a finite set $\{x_1, \ldots, x_n\}$, and $|\mu| = \sum_{x_i \in X} \mu(x_i)$, the crisp cardinality of μ, then the function $m(\mu) = \frac{|\mu|}{n}$, is a fuzzy measure.

Remember, that since $\mu \cdot \sigma = T \circ (\mu \times \sigma)$, $\mu + \sigma = S \circ (\mu \times \sigma)$, it is

$$\mu \cdot \sigma \leqslant \mu, \mu \cdot \sigma \leqslant \sigma, \mu \leqslant \mu + \sigma, \sigma \leqslant \mu + \sigma.$$

Provided $\mu \cdot \sigma$, $\mu + \sigma \in \mathfrak{F}$, for all fuzzy measure m, is:

$$m(\mu \cdot \sigma) \leqslant m(\mu), m(\mu \cdot \sigma) \leqslant m(\sigma), m(\mu) \leqslant m(\mu + \sigma), m(\sigma) \leqslant m(\mu + \sigma),$$

and

$$m(\mu \cdot \sigma) \leqslant min(m(\mu),\ m(\sigma)) \leqslant max(m(\mu),\ m(\sigma)) \leqslant m(\mu + \sigma).$$

In the particular case in which $\mu, \sigma \in \{0, 1\}^X$, it results

$$m(A \cap B) \leqslant min(m(A),\ m(B)) \leqslant max(m(A),\ m(B)) \leqslant m(A \cup B),$$

for all $A, B \in \mathbb{P}(X)$. Notice that fuzzy measures can be applied to both $[0, 1]^X$ and $\mathbb{P}(X)$.

7.3 Types of Measures

Given a triplet $(X, \mathfrak{F}, m)$, where m is a fuzzy measure, if for some negation $'(N)$ and some union $+(S)$, is

1. When $\mu \leqslant \sigma'$, then $m(\mu + \sigma) \leqslant m(\mu) + m(\sigma)$, m is sub-additive
2. When $\mu \leqslant \sigma'$, then $m(\mu + \sigma) \geqslant m(\mu) + m(\sigma)$, m is super-additive,

and when m is both sub-additive and super-additive, that is

$$\text{When } \mu \leqslant \sigma', \quad \text{then} \quad m(\mu + \sigma) = m(\mu) + m(\sigma), \quad m \text{ is additive.}$$

This classification (once completed with those measures that are neither sub-additive, nor super-additive), in the case in which $\mu, \sigma \in \{0, 1\}^X$, particularizes to:

- If $A \cap B = \varnothing$, and $m(A \cup B) \leqslant m(A) + m(B)$, m is sub-additive
- If $A \cap B = \varnothing$, and $m(A \cup B) \geqslant m(A) + m(B)$, m is super-additive
- If $A \cap B = \varnothing$, and $m(A \cup B) = m(A) + m(B)$, m is additive.

Example 7.3.1 The measure $m(A) = \frac{|A|}{n}$, in a finite set $X = \{x_1, \ldots, x_n\}$, is additive.

7.4 λ-Measures

With $\mathfrak{F} = \mathbb{P}(X)$, $m_\lambda : \mathbb{P}(X) \to [0, 1]$ is called a Sugeno's λ-*measure* if, with $\lambda > -1$, it is:

1. $m_\lambda(\varnothing) = 0$
2. $m_\lambda(X) = 1$
3. If $A \cap B = \varnothing$, $m_\lambda(A \cup B) = m_\lambda(A) + m_\lambda(B) + \lambda m_\lambda(A) m_\lambda(B)$.

Theorem 7.4.1 *All mapping m_λ is, actually, a fuzzy measure.*

Proof What lacks to be proven is that $A \subset B$ implies $m_\lambda(A) \leqslant m_\lambda(B)$. Since $A \subset B \Leftrightarrow B = A \cup (A^C \cap B)$, with $A \cap (A^C \cap B) = \varnothing$, it follows $m_\lambda(B) = m_\lambda(A^C \cap B) + \lambda m_\lambda(A) m_\lambda(A^C \cap B) = m_\lambda(A)[1 + \lambda m_\lambda(A^C \cap B)] + m_\lambda(A^C \cap B)$. From $\lambda > -1$, it follows $1 + \lambda m_\lambda(A^C \cap B) > 1 - m_\lambda(A^C \cap B)$, and $m_\lambda(B) > m_\lambda(A)[1 - m_\lambda(A^C \cap B)] = m_\lambda(A) - m_\lambda(A) m_\lambda(A^C \cap B) > m_\lambda(A)$. □

Theorem 7.4.2 $m_\lambda(A^C) = \frac{1 - m_\lambda(A)}{1 + \lambda m_\lambda(A)}$, *for all $A \in \mathbb{P}(X)$.*

Proof From $A \cap A^C = \varnothing$, follows $m_\lambda(X) = 1 = m_\lambda(A \cup A^C) = m_\lambda(A) + m_\lambda(A^C) + \lambda m_\lambda(A) m_\lambda(A^C) = m_\lambda(A^C)[1 + \lambda m_\lambda(A)] + m_\lambda(A)$. □

Notice, that it is $m_\lambda(A^C) = N_\lambda(m_\lambda(A))$, with the Sugeno's negation $N_\lambda(x) = \frac{1-x}{1+\lambda x}(\lambda > -1)$.

Remarks 7.4.3 • With the t-conorm $S_\lambda(x, y) = x+y+\lambda xy$, it follows $m_\lambda(A \cup B) = S_\lambda(m_\lambda(A), m_\lambda(B))$, if $A \cap B = \varnothing$.

- When $\lambda = 0$, m_0 is just a probability defined in $\mathbb{P}(X)$ since it results $m_\lambda(A \cup B) = m_\lambda(A) + m_\lambda(B)$ when $A \cap B = \varnothing$, that is, m_0 is an additive measure. In addition, since $N_0(x) = 1 - x$, it is $m_0(A^C) = 1 - m_0(A)$.
- Notice that the axioms required for a λ-measure do not individuate a single one of them. For example, with $\lambda = 0$ what is obtained is the set of all probabilities on $\mathbb{P}(X)$.
- If $\lambda \in (-1, 0)$, it results

$$A \cap B = \varnothing : m_\lambda(A \cup B) \leqslant m_\lambda(A) + m_\lambda(B),$$

  that is, *if* $\lambda \in (-1, 0)$, *all the corresponding* λ*-measures are sub-additive*. As it is easy to prove, if $\lambda \in (0, +\infty)$, m_λ is super-additive.
- As it is well known, if X is a finite set $X = \{x_1, \ldots, x_n\}$, all probabilities $m_0 : \mathbb{P}(X) \to [0, 1]$ are defined by choosing n numbers $m_0(\{x_i\}) \in [0, 1]$, $1 \leqslant i \leqslant n$, verifying $\sum_{i=1}^{n} m_0(\{x_i\}) = 1$, because $1 = m_0(\{x_1, \ldots, x_n\}) = m_0(x_1) + \ldots + m_0(x_n)$. Something analogous happens with λ-measures when $X = \{x_1, \ldots, x_n\}$. For example, if $X = \{x_1, x_2, x_3\}$, it follows

$$\begin{aligned} 1 &= m_\lambda(X) = m_\lambda(\{x_1, x_2, x_3\}) \\ &= m_\lambda(\{x_1, x_2\} \cup \{x_3\}) = m_\lambda(\{x_1, x_2\}) + m_\lambda(x_3) + \lambda m_\lambda(\{x_1, x_2\}) m_\lambda(x_3) \\ &= \sum_{i=1}^{3} m_\lambda(x_i) + \lambda \sum_{1=i<j=3} m_\lambda(x_i) m_\lambda(x_j) + \lambda^2 m_\lambda(x_1) m_\lambda(x_2) m_\lambda(x_3). \end{aligned}$$

  and for each $\lambda \in (-1, +\infty)$, the values $m_\lambda(x_i)$, $1 \leqslant i \leqslant 3$, are to be taken verifying this equation that, of course, with $\lambda = 0$ reduces to $1 = \sum_{i=1}^{3} m_0(x_i)$.

  In the case that, for example, is $m(x_1) = 0$, follows

$$1 = m_\lambda(x_2) + m_\lambda(x_3) + \lambda[m_\lambda(x_2) m_\lambda(x_3)] = m_\lambda(x_2) + m_\lambda(x_3) + \lambda m_\lambda(x_2) m_\lambda(x_3)$$

  that is: $m_\lambda(x_3) = \frac{1-m_\lambda(x_2)}{1+\lambda m_\lambda(x_2)}$. With $m_\lambda(x_2) = 0.7$, results $m_\lambda(x_3) = \frac{0.3}{1+0.3\lambda}$. With $\lambda = 1$: $m_1(x_3) = \frac{0.3}{1.03} = 0.29$. That is: a measure m_1 is defined in $X = \{x_1, x_2, x_3\}$, by $m_1(x_1) = 0$, $m_1(x_2) = 0.7$, $m_1(x_3) = 0.29$. Notice that, since m_1 is not a probability, it is $\sum_{i=1}^{3} m(x_i) = 0.99 < 1$.

7.5 Measures of Possibility and Necessity

Let it be $\mathfrak{F} \subset \mathbb{P}(X)$ a Boolean algebra of subsets of X. A mapping $\pi : \mathfrak{F} \to [0, 1]$ is called a *measure of possibility*, provided:

- $\pi(\varnothing) = \varnothing$
- $\pi(X) = 1$
- $\pi(A \cup B) = max(\pi(A), \pi(B))$, for all A, B in $\mathfrak{F}$.

Notice that the last property does not require $A \cap B = \varnothing$. Actually, any of these mappings are fuzzy measures, since:

$$A \subset B \Leftrightarrow A \cup B = B\text{: } \pi(B) = max(\pi(A), \pi(B)) \geqslant \pi(A), \text{ or } \pi(A) \leqslant \pi(B).$$

From $max(\pi(A), \pi(B)) \leqslant \pi(A) + \pi(B)$, it follows $\pi(A \cup B) \leqslant \pi(A) + \pi(B)$ even if $A \cap B = \varnothing$. Hence, possibility measures are sub-additive.

From $A \cup A^C = X$, it is $1 = max(\pi(A), \pi(A^C)) \leqslant \pi(A) + \pi(A^C)$, or $1 - \pi(A) \leqslant \pi(A^C)$.

Obviously,

$$\pi(A_1 \cup A_2 \cup \ldots \cup A_n) = max(\pi(A_1), \pi(A_2), \ldots, \pi(A_n)), \quad \text{for all} \quad A_1, \ldots, A_n \text{ in } \mathfrak{F}.$$

Hence, if $X = \{x_1, \ldots, x_n\}$ is a finite set, to have a possibility measure π, its values $\pi(x_i)$ do verify:

$$1 = \pi(X) = max(\pi(x_1), \pi(x_2), \ldots, \pi(x_n)),$$

forcing that some of the values $\pi(x_i)$ should equal 1. For example, if $X = \{x_1, x_2, x_3\}$, the three values $\pi(x_1) = 0$, $\pi(x_2) = 0.5$, $\pi(x_3) = 1$, define a particular measure of possibility on $\mathbb{P}(X)$. It is, for example, $\pi(\{x_1, x_2\}) = max(0, 0.5) = 0.5$, $\pi(\{x_1, x_3\}) = max(0, 1) = 1$, etc.

Remark 7.5.1 Instead of a family $\mathfrak{F}$ of crisp sets no problem arises in considering a family of fuzzy sets. Measures of possibility can be applied to fuzzy sets with the only changes of $A \cap B = \varnothing$ by $\mu \leqslant \sigma'$, and $A \cap B \neq \varnothing$ by $\mu \nleqslant \sigma'$. The only caution is to use the connectives min, max to preserve distributivity.

Theorem 7.5.2 *For each $\mu \in [0, 1]^X$ such that $Sup\, \mu = 1$, the mapping $\pi_\mu : \mathfrak{F} \to [0, 1]$ defined by*

$$\pi_\mu(A) = \underset{x \in X}{Sup}\; min(\mu(x), \mu_A(x)),\ A \in \mathfrak{F},$$

is a possibility measure.

Proof $\pi_\mu(\varnothing) = \underset{x \in X}{Sup}\; min(\mu(x), 0) = 0$. $\pi_\mu(X) = \underset{x \in X}{Sup}\; min(\mu(x), 1) = \underset{x \in X}{Sup}\; \mu(x) = 1$. Finally, since $\mu_{A \cup B}(x) = max(\mu_A(x), \mu_B(x))$ for all $x \in X$:

$$\begin{aligned}\pi_\mu(A \cup B) &= \underset{x\in X}{Sup}\, min(\mu(x), \mu_{A\cup B}(x)) = \underset{x\in X}{Sup}\, min(\mu(x), max(\mu_A(x), \mu_B(x)))\\ &= \underset{x\in X}{Sup}\, max(min(\mu(x), \mu_A(x)), min(\mu(x), \mu_B(x)))\\ &= max(\underset{x\in X}{Sup}\, min(\mu(x), \mu_A(x)), \underset{x\in X}{Sup}\, min(\mu(x), \mu_B(x)))\\ &= max(\pi_\mu(A), \pi_\mu(B)).\end{aligned}$$ □

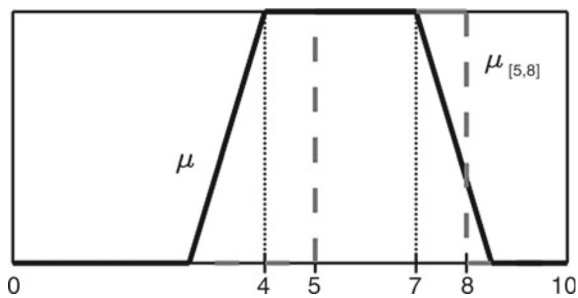
It is said that μ is the *possibility distribution* of π_μ, a possibility measure that can be considered as the one "conditioned" by μ. Thus, each fuzzy set $\mu \in [0, 1]^X$ such that $Sup\, \mu = 1$ (μ is never self-contradictory) induces the possibility measure π_μ. Provided $X = \{x_1, \ldots, x_n\}$ is a finite set, it results

$$\pi_\mu(A) = \underset{1\leqslant i\leqslant n}{Max}\; min(\mu(x_i), \mu_A(x_i)).$$

Hence, all non self-contradictory *fuzzy sets*, those μ such that $Sup\, \mu = 1$, *can be viewed as possibility distributions*.

Examples 7.5.3 1. If $X = \{x_1, x_2, x_3\}$ and $\mu = 1|x_1 + 0.6|x_2 + 0.8|x_3$, follows $\pi_\mu(X) = Max(min(1, \mu_A(x_1)), min(0.6, \mu_A(x_2)), min(0.8, \mu_A(x_3))$ that, with $A = \{x_2, x_3\}$ gives $\pi_\mu(A) = max(0, 0.6, 0.8) = 0.8$, and with $A = \{x_1, x_3\}$ gives $\pi_\mu(A) = max(min(1, 1), min(0.6, 0), min(0.8, 1)) = max(1, 0, 0.8) = 1$.

2. If $X = [0, 10]$, $A = [5, 8]$, with μ in the next figure, it is $\pi_\mu([5, 8]) = \underset{x\in[0,10]}{Sup}\; min(\mu(x), \mu_{[5,18]}(x)) = \underset{x\in[0,10]}{Sup} \begin{Bmatrix} 0, & x \in [0, 5] \cup [8, 10] \\ \mu(x), & x \in [5, 18] \end{Bmatrix} = 1.$



A mapping $N : \mathfrak{F} \to [0, 1]$ is a *measure of necessity* provided

- $N(\varnothing) = 0$
- $N(X) = 1$
- $N(A \cap B) = min(N(A), N(B))$, for all A, B in $\mathfrak{F}$.

Since $A \subset B \Leftrightarrow A \cap B = A$, it is $N(A) = min(N(A), N(B)) \leqslant N(B)$. Hence, all these mappings are actually fuzzy measures. From $min(N(A), N(B)) \leqslant \pi(A)+\pi(B)$, it follows $N(A \cap B) \leqslant \pi(A)+\pi(B)$, Then $0 = N(\varnothing) = N(A \cap A^C) = min(N(A), N(A^C))$, that implies:

$$\text{Either } N(A) = 0, \text{ or } N(A^C) = 0, \text{ and } N(A) + N(A^C) \leqslant 1, \text{ or}$$
$$N(A^C) \leqslant 1 - N(A).$$

Obviously, if $A_1, A_2, \ldots, A_n \in \mathfrak{F} : N(A_1 \cap A_2 \cap \ldots \cap A_n) = min(N(A_1), N(A_2), \ldots, N(A_n))$. Then if $X = \{x_1, \ldots, x_n\}$, to define a necessity measure it is needed to take $N(x_1), N(x_2), \ldots, N(x_3)$ such that: $0 = min(N(x_1), N(x_2), \ldots, N(x_n))$, an equality that forces some of the values $N(x_i)$ to be 0.

Theorem 7.5.4 *Given a possibility measure* $\pi : \mathfrak{F} \to [0, 1]$*, the function* $N_\pi(A) = 1 - \pi(A^C)$*, for all* $A \in \mathfrak{F}$*, is a necessity measure.*

Proof $N_\pi(\varnothing) = 1 - \pi(X) = 0$. $N_\pi(X) = 1 - \pi(\varnothing) = 1$. $N_\pi(A \cap B) = 1 - \pi(A^C \cup B^C) = 1 - max(\pi(A^C), \pi(B^C)) = min(1 - \pi(A^C), 1 - \pi(B^C)) = min(N_\pi(A), N_\pi(B))$. □

Theorem 7.5.5 *Given a necessity measure* $N : \mathfrak{F} \to [0, 1]$*, the function* $\pi_N(A) = 1 - N(A^C)$*, for all* $A \in \mathfrak{F}$ *is a possibility measure.*

Proof $\pi_N(\varnothing) = 1 - N(X) = 0$. $\pi_N(X) = 1 - N(\varnothing) = 1$. $\pi_N(A \cup B) = 1 - N(A^C \cap B^C) = 1 - min(N(A^C), N(B^C)) = max(\pi_N(A), \pi_N(B))$. □

The pairs (π, N_π) and (N, π_N) are called dual-pairs of possibility/necessity measures. Notice that with them it can be read:

$$Necessity\ of\ A = \text{“}not\ the\ possibility\ of\ not\ A\text{”}$$
$$Possibility\ of\ A = \text{“}not\ the\ necessity\ of\ not\ A\text{”}.$$

Remark 7.5.6 If $\pi = \pi_\mu$, the corresponding N_{π_μ} is given by

$$N_{\pi_\mu}(A) = 1-\pi_\mu(A) = 1-\underset{x\in X}{Sup}\ min(\mu(x), \mu_A(x)) = \underset{x\in X}{Inf}\ max(1-\mu(x), \mu_{A^C}(x)),$$

that, in the finite case $X = \{x_1, \ldots, x_n\}$, is

$$N_{\pi_\mu}(A) = \min_{1\leqslant i\leqslant n} max(1 - \mu(x_i), \mu_{A^C}(x_i)).$$

Theorem 7.5.7 *For all dual pair* (π, N) *is:*

1. *If* $N(A) > 0$*, then* $\pi(A) = 1$
2. *If* $\pi(A) < 1$*, then* $N(A) = 0$

Proof

$$1 = max(\pi(A), \pi(A^C)), \quad \text{and} \quad 0 = min(N(A), N(A^C)),$$

follows that if $N(A) > 0$, then $N(A^C) = 0$, and $\pi(A) = 1 - N(A^C) = 1$. If $\pi(A) < 1$, then $\pi(A^C) = 1$ and $N(A) = 1 - \pi(A^C) = 0$ □

Remark 7.5.8 Although the proof will not be presented, let's show the following important notice. *In the case X is finite, for any possibility measure π it exists a (non unique!) fuzzy set $\mu \in [0, 1]^X$ with Sup $\mu = 1$ such that $\pi = \pi_\mu$. In the finite case, all possibility measures come from possibility distributions.*

7.6 Examples

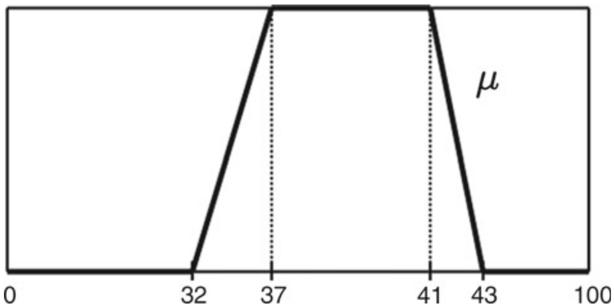
Example 7.6.1 On the age of a person p it is only available the incomplete information given by

1. $37 \leqslant Age(p) \leqslant 41$
2. It is neither $Age(p) \leqslant 32$, nor $Age(p) \geqslant 43$.

What can be said about the possibility and the necessity of "$Age(p) \geqslant 42$", "$Age(p) \leqslant 40$", and "$Age(p) \geqslant 33$"?

Solution

The available incomplete information can be represented by the following possibility distribution μ:



Hence,

- $\pi_\mu(Age(p) \geqslant 42) = \pi_\mu([42, 100]) = \underset{x \in [0,100]}{Sup}\ min(\mu(x), \mu_{[42,100]}(x)) = \underset{x \in [42,100]}{Sup}\ \mu(x) = \mu(42) : \in (0, 1)$. Hence $N_{\pi_\mu}(Age(p) \geqslant 42) = 0$.

  The value $\mu(42)$ can be computed as follows. The segment between $(41, 1)$ and $(43, 0)$, verifies

$$0 = \begin{vmatrix} x & y & 1 \\ 41 & 1 & 1 \\ 43 & 0 & 1 \end{vmatrix} = x + 2y - 43 \Rightarrow y = \frac{43 - x}{2},$$

hence, $\mu(42) = \frac{43-42}{2} = 0.5$. The possibility of "$Age(p) \geqslant 42$" is 0.5, and its necessity is 0.

- $\pi_\mu(Age(p) \leqslant 40) = \pi_\mu([0, 40]) = \mu(40) = 1$, and $N_{\pi_\mu}(Age(p) \leqslant 40) = N_{\pi_\mu}([0, 40]) = 1 - \pi_\mu((40, 100]) = 1 - \underset{x \in (40,100]}{Sup} \mu(x) = 1 - \mu(40) = 0$. The possibility of "$Age(p) \geqslant 42$" is 1, and its necessity is 0.
- $\pi_\mu(Age(p) \geqslant 33) = \pi_\mu([33, 100]) = 1$, and $N_{\pi_\mu}(Age(p) \geqslant 33) = N_{\pi_\mu}([33, 100]) = 1 - \pi_\mu([0, 33)) = 1 - \mu(33)$. This value can be computed by:

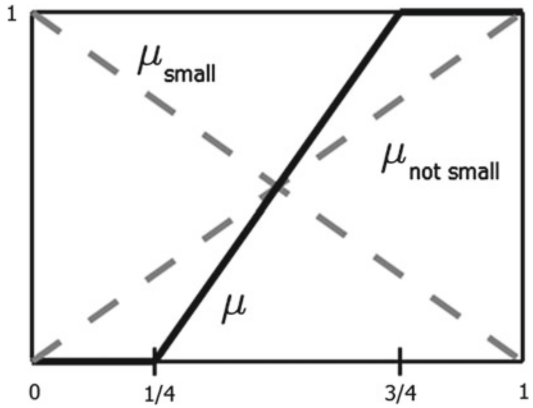
$$0 = \begin{vmatrix} x & y & 1 \\ 32 & 0 & 1 \\ 37 & 1 & 1 \end{vmatrix} = -x + 5y + 32 \Rightarrow y = \frac{x - 32}{5}, \quad \text{and} \quad \mu(33) = \frac{1}{5}.$$

That is, $N_{\pi_\mu}(Age(p) \geqslant 33) = 1 - \frac{1}{5} = \frac{4}{5}$. The possibility of "$Age(p) \geqslant 33$" is 1, and its necessity is 0.8.

Example 7.6.2 It is only known that a function $F : X \to [0, 1]$ doesn't take any value below $1/4$ but takes values above $3/4$. What can be said on the possibility and necessity of the imprecise statements "*F is small*", and "*F is not small*"?

Solution

The available but incomplete information can be translated by the possibility distribution μ



Let's take $\mu_{small}(x) = 1 - x$, $\mu_{not\ small}(x) = x$. It is:

- $\pi_\mu(F\ is\ small) = \underset{x\in[0,1]}{Sup}\ min(\mu(x), \mu_{small}(x)) = \underset{x\in[0,1]}{Sup}\ min(\mu(x), 1 - x) = 1/2 \Rightarrow N_{\pi_\mu}(F\ is\ small) = 1 - \pi_\mu(F\ is\ not\ small)$
- $\pi_\mu(F\ is\ not\ small) = \underset{x\in[0,1]}{Sup}\ min(\mu(x), \mu_{not\ small}(x)) = \underset{x\in[0,1]}{Sup}\ min(\mu(x), x) = 1 \Rightarrow N_{\pi_\mu}(F\ is\ small) = 1 - \pi_\mu(F\ is\ small).$

  Hence:

  - $\pi_\mu(F\ is\ small) = 0.5$, and $N_{\pi_\mu}(F\ is\ small) = 1 - 1 = 0.$
  - $\pi_\mu(F\ is\ not\ small) = 1$, and $N_{\pi_\mu}(F\ is\ not\ small) = 1 - 0.5 = 0.5.$

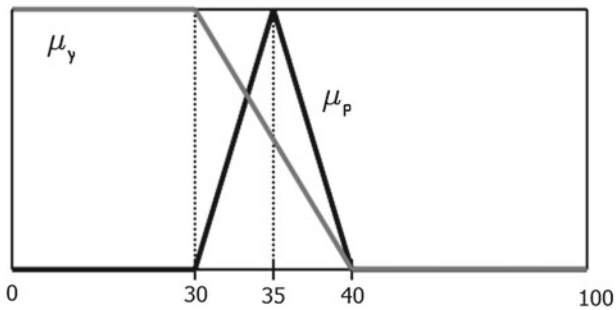
Remark 7.6.3 Notice that this example is, like the following, not with questions related to precise or crisp sets, but to impreciseness (fuzzy sets). Although Possibility Theory is introduced with crisp sets, it is also applicable to fuzzy sets within the theory given by the triplet $(min, max, 1 - id)$.

Example 7.6.4 John is a member of a community where the predicate $Y = young$ is used following

$$\mu_Y(x) = \begin{cases} 1, & \text{if } x \leqslant 30 \\ 0, & \text{if } x > 40 \\ \frac{40-x}{10}, & \text{if } 30 < x \leqslant 40 \end{cases}$$

Find the possibility and the necessity of the statement "John is around 35 years old".

Solution. The graphics of μ_Y and μ_P, with $P = Around\ 35$, are



Hence, π_μ(John is around 35 years old) $= Sup_{x\in[0,100]}\, min(\mu_Y(x), \mu_P(x)) = \frac{2}{3}$, since it does correspond with the intersection of μ_Y and μ_P, that is, of the straight lines respectively given by $y = \frac{40-x}{10}$, $y = \frac{x-3}{5}$. It results $x = \frac{100}{3}$ and $y = \frac{2}{3}$. Since π_μ(John is around 35 years old) < 1, it results N_{π_μ}(John is around 35 yearsold) $= 0$.

7.7 Probability, Possibility and Necessity

In the case where $\mathfrak{F}$ is a Boolean sub-algebra of $\mathbb{P}(X)$, also probabilities $p : \mathfrak{F} \to [0, 1]$ can be taken into account. Once a dual-pair (N, π) is given, it appears the problem of which probabilities are *consistent* with (N, π).

The property $N(A) + N(A^C) \leqslant 1$, equivalent to $N(A) \leqslant 1 - N(A^C) = \pi(A)$, shows that for all $A \in \mathfrak{F}$ it is

$$N(A) \leqslant \pi(A),$$

provided the pair (N, π) is dual. A probability p is said *consistent* with the dual pair (N, π) if

$$N(A) \leqslant p(A) \leqslant \pi(A), \quad \text{for all} \quad A \in \mathfrak{F}.$$

In this hypothesis, if $N(A) > 0$, it is also $p(A) > 0$, and $\pi(A) > 0$, and if $\pi(A) = 0$ it is $N(A) = p(A) = 0$. That is,

- If something is just a little bit necessary, it is probable and possible.
- If something is not possible at all, it is neither necessary nor probable.

Analogously, if $p(A) > 0$ it is $\pi(A) > 0$ although it could be $N(A) = 0$. That is,

- If something is just a little bit probable, it is possible, but not necessarily necessary.

Notice that from $N(A) \leqslant p(A) \leqslant \pi(A)$, or equivalently from $1 - \pi(A^C) \leqslant p(A) \leqslant \pi(A)$, follows $1 - \pi(A) \leqslant p(A^C) \leqslant \pi(A^C)$.

Example 7.7.1 Let is $X = \{1, 2, 3\}$, and $\mu = 0.7|1 + 1|2 + 0.5|3$. Then, with $\pi_\mu(A) = \underset{i \in X}{Sup}\, min(\mu(i), \mu_A(i))$, and $N_\mu^{(A)} = 1 - \pi_\mu(A)$, we get:

1. $\pi_\mu(1) = \mu(1) = 0.7$, $\pi_\mu(2) = \mu(2) = 1$, $\pi_\mu(3) = \mu(3) = 0.5$.
   $\pi_\mu(\{1, 2\}) = max(\mu(1), \mu(2)) = 1$, $\pi_\mu(\{1, 3\}) = max(0.7, 0.5) = 0.7$, $\pi_\mu(\{2, 3\}) = 1$, $\pi_\mu\{X\} = 1$.
2. $N_\mu(1) = 1 - \pi_\mu(\{2, 3\}) = 1 - 1 = 0$, $N_\mu(2) = 1 - \pi_\mu(\{1, 3\}) = 1 - 0.7 = 0.3$, $N_\mu(3) = 1 - \pi_\mu(\{1, 2\}) = 0$
   $N_\mu(\{1, 2\}) = 1 - \pi_\mu(3) = 0.5$, $N_\mu(\{1, 3\}) = 1 - \pi_\mu(2) = 0$, $N_\mu(\{2, 3\}) = 1 - \pi_\mu(1) = 0.3$, $N_\mu(X) = 1$.

Hence, the consistent probabilities are given by triplets $p(1)$, $p(2)$, $p(3)$ in $[0, 1]$ such that $p(1) + p(2) + p(3) = 1$, and verifying:

$$0 \leqslant p(1) \leqslant 0.7, 0.3 \leqslant p(2) \leqslant 1, 0 \leqslant p(3) \leqslant 0.5.$$

For instance, with $p(1) = 0.4$, $p(2) = 0.5$, $p(3) = 0.1$, we have a consistent probability, as well as with $p(1) = 0.5$, $p(2) = 0.3$, $p(3) = 0.2$. But the probability given by $p(1) = 0.6$, $p(2) = 0.2$, $p(3) = 0.2$ is not consistent, because of $p(2) < 0.3$. With it there is an element whose probability is smaller than its necessity. In the same vein, the probability given by the triplet $p(1) = 0.1$, $p(2) = 0.3$, $p(3) = 0.6$, is also non-consistent because one of the probabilities is greater than the corresponding possibility.

7.8 Probability of Fuzzy Sets

Let's shortly formalize the classical concept of probability. In a universe X, let $\mathfrak{F} \subset \mathbb{P}(X)$ be a Boolean algebra of parts of X, that is,

- If $A, B \in \mathfrak{F} \Rightarrow A \cap B, A \cup B, A^C, B^{\subset} \in \mathfrak{F}$.
- $\varnothing \in \mathfrak{F}, X \in \mathfrak{F}$.

It is said that $p : \mathfrak{F} \to [0, 1]$ is a probability in $(X, \mathfrak{F})$, provided

- $p(\varnothing) = 0$
- If $A \cap B = 0$, then $p(A \cup B) = p(A) + p(B)$.

Theorem 7.8.1 *1.* $p(A^C) = 1 - p(A)$*, for all* $A \in \mathfrak{F}$*.*
2. If $A \subset B$*, then* $p(A) \leqslant p(B)$ *(that is p is a measure)*
3. $p(X) = 1$.
4. $p(A \cup B) + p(A \cap B) = p(A) + p(B)$*, for all* $A, B \in \mathfrak{F}$*.*

Proof Items (1) and (2) just follow from the fact that p is a 0-measure. $p(X) = p(\varnothing^{\subset}) = 1 - p(\varnothing) = 1$. Finally, since

$$A \cup B = (A \cap B) \cup (B - A) \cup (A - B),$$

with $(A \cap B) \cap (B - A) = \varnothing$, $(A \cap B) \cap (A - B) = \varnothing$, and $(B - A) \cap (A - B) = \varnothing$, follows

$$p(A \cup B) = p(A \cap B) + p(B - A) + p(A - B).$$

But, from $A = (A - B) \cup (A \cap B)$, and $B = (B - A) \cup (A \cap B)$, it also follows (since the unions are disjunct):

- $p(A) = p(A - B) + p(A \cap B) \Rightarrow p(A - B) = p(A) - p(A \cap B)$
- $p(B) = p(B - A) + p(A \cap B) \Rightarrow p(B - B) = p(B) - p(A \cap B)$.

Hence, $p(A \cup B) = p(A \cap B) + p(A) - p(A \cap B) + p(B) - p(A \cap B) = p(A) + p(B) - p(A \cap B)$. $\square$

Of special importance for the applications are the probabilities defined in the real line $\mathbb{R}$, with $\mathfrak{F} = \mathfrak{B}$ the so-called *Borel's algebra*, given by all the unions, complements and intersections of the open, closed, semi-open, and semi-closed intervals of $\mathbb{R}$. Then, if $A \in \mathfrak{B}$, the probability of A is *defined* by the Lebesgue-Stieltjes integral

$$p(A) = \int_A dP = \int_R \mu_A(x)dx = E(\mu_A),$$

that is, as the mathematical expectation of μ_A. Then, if $\mu \in [0, 1]^{\mathbb{R}}$ is Borel-measurable, it can be analogously defined

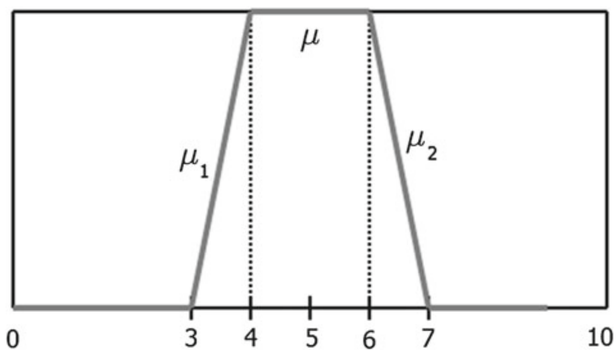
$$p(\mu) = E(\mu) = \int_R \mu(d)dx$$

Obviously, $p(\mu_0) = E(\mu_0) = 0$, $p(\mu_1) = E(\mu_1) = 1$, and if $\mu \leqslant \sigma$ follows $p(\mu) \leqslant p(\sigma)$. In addition, with $\mu \cdot \sigma = min \circ (\mu \times \sigma)$, $\mu + \sigma = max \circ (\mu \times \sigma)$, it is

$$p(\mu + \sigma) + p(\mu \cdot \sigma) = p(\mu) + p(\sigma),$$

that implies: If $\mu \cdot \sigma = \mu_0$, then $p(\mu + \sigma) = p(\mu) + p(\sigma)$.

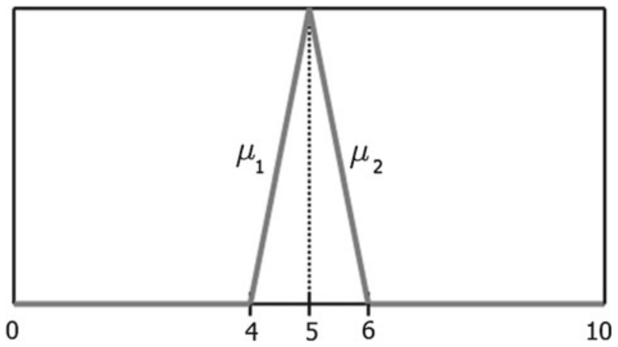
Example 7.8.2 Which is the probability of the fuzzy set μ (fuzzy event) given by



Solution

$10 \cdot p(\mu) = \int_{[0,10]} \mu dx = \int_{[3,7]} \mu dx = \int_{[3,4]} \mu_1 dx + \int_{[4,6]} 1dx + \int_{[6,7]} \mu_2 dx = 2 + \int_{[3,4]}(x-3)dx + \int_{[6,7]}(7-x)dx = 2 + \frac{1}{2} + \frac{1}{2} = 3$. Then $p(\mu) = \frac{3}{10} = 0.3$.

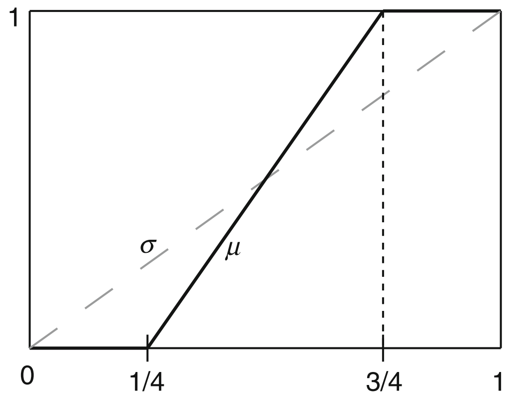
Example 7.8.3 Which is the probability of the fuzzy event μ



Solution

$10 \cdot p(\mu) = \int_{[0,10]} \mu dx = \int_{[4,5]} \mu_1 dx + \int_{[5,6]} \mu_2 dx = \int_4^5 (4-x)dx + \int_5^6 (6-x)dx = -\frac{1}{2} + \frac{3}{2} = 1$. Then $p(\mu) = 0.1$.

Example 7.8.4 Which is the probability of the fuzzy event μ:

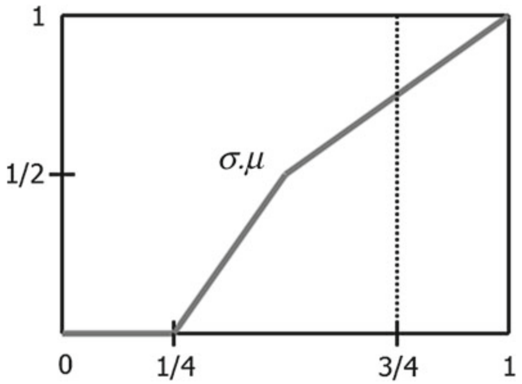


Which is the conditional probability $p(\sigma/\mu)$?

Solution

$p(\mu) = \int_{[0,10]} \mu dx = \int_{1/4}^{3/4} \mu_1 dx + \int_{3/4}^{1} 1 dx = \frac{1}{4} + \int_{1/4}^{3/4} (2x - \frac{1}{2})dx = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$.

$p(\sigma|\mu) = \frac{p(\sigma \cdot \mu)}{p(\mu)} = 2p(\sigma \cdot \mu)$, with $\sigma \cdot \mu = \min \circ (\sigma \times \mu)$, given by



Then $p(\sigma \cdot \mu) = \int_{1/4}^{1/2} (4 - x)dx + \int_{1/2}^{1} x dx = \frac{7}{16}$, and $p(\sigma|\mu) = \frac{7/16}{1/2} = \frac{7}{8}$.

Chapter 8
An Introduction to Fuzzy Control

8.1 Introduction

Probably one of the most successful developments of fuzzy reasoning, from the industrial point of view, is the design of fuzzy control systems, also called *linguistic* control systems, or simpler, the applications of fuzzy controllers.

A fuzzy control system is based on a set of fuzzy "if-then" rules of behavior that consider the kind of stimuli from the environment, that the system will receive, meanwhile at a given time, the values of these stimuli represent the facts, that the rules have to consider to offer proper actions.

As it has been proved from its origins, fuzzy control should be useful in situations where (a) There is no acceptable mathematical model for the plant, (b) There are experienced human operators who can satisfactorily control the plant and provide qualitative control rules in terms of vague and fuzzy sentences, and (c) In applications where there is a large uncertainty or unknown variation in plant parameters and structures. In this cases, fuzzy control can be considered as a model-free approach and it does not require a mathematical model of the objective plant. It is referred to as a knowledge-based control approach, and it makes an effective use of all available information related to the system, from sensors which provide numerical measurements of key variables to human experts who provide linguistic descriptions about the system and control instructions.

However, currently fuzzy control approach is mainly devoted to model-based methods. On the one hand, there are cases in which operators cannot precisely tell their action in a particular situation, or simply, operator's control may not be always optimal with respect to some performance objective. In this sense, identification for obtaining fuzzy models from process data is very important.

On the other hand, although the uncertainty in dynamic systems can be handled using appropriate control actions (e.g., their sensitivity to external disturbances and parameter changes can be reduced stabilizing an unstable system, or the systems dynamic behavior could be modified speeding up a slow system), the application of control action over a system can potentially destabilize stable plants. Thus,

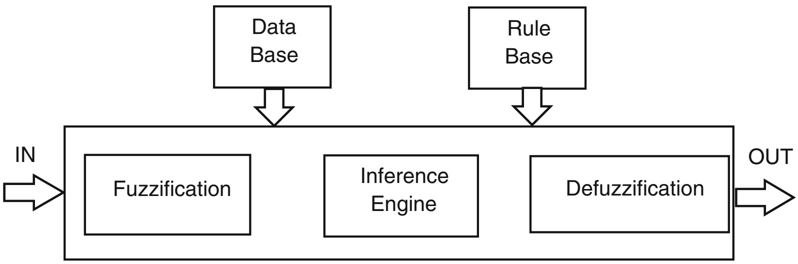
E. Trillas and L. Eciolaza, *Fuzzy Logic*, Studies in Fuzziness and Soft Computing 320, DOI 10.1007/978-3-319-14203-6_8

stability analysis is a major issue in control design. A model-based fuzzy control approach allows performing system stability, performance and robustness analysis. However, this approach also implies the use of mathematically involved and rather non-transparent techniques to ensure robust performance.

When dealing with the control of nonlinear systems, a conventional approach to design controllers is the derivation of fuzzy models from given nonlinear system equations. Dynamic systems are modeled in the state space framework, using a state transition model, which describes the evolution of the states over time. Fuzzy logic provides a simple and straightforward way to decompose the task of modeling and control design into a group of local tasks, which tend to be easy to handle. In the end, fuzzy logic also provides the mechanism to blend these local tasks together to deliver the overall model and control design.

Advances in modern control have made available a large number of powerful design tools, specially in the case of linear control systems. These tools range from state space optimal control to the robust control paradigms. In the case of Takagi-Sugeno (TS) fuzzy models, a local linear system description is used for each rule, which then uses a control methodology to fully take advantage of the advances of modern control theory. From a theoretical point of view, model-based fuzzy control can meet established rigor and systematic control synthesis of conventional control theory in terms of stability analysis, systematic design procedures, incorporation of performance specifications, robustness and optimality.

A fuzzy controller has the general structure shown in the following figure. Three main blocks may be distinguished: a data base, a rule base and a processing unit.



The data base contains all information needed to specify a particular configuration of the fuzzy controller. It will have information related to the number, shape and distribution of the fuzzy sets specifying the meaning of the linguistic terms of each linguistic variable involved in the application. The data base also contains information related to the kind of crisp-fuzzy and fuzzy-crisp conversions, and the type of operations for numerical calculations of the conjunction or disjunction premises. Similarly with respect to the availability of operations for the implication, that realizes the "if-then" connective of the rules, as well as for the operation that computes the aggregation of the conclusions of several rules that might be simultaneously activated by the prevailing conditions of the environment.

The rule base contains the set of rules that will govern the behavior of the controller. Through the rule base the task of modeling and control design is decomposed into a group of local tasks, which tend to be easy to handle.

The processing unit cares for the compatibility of data (input and output interfaces) and for the execution of the rules (inference engine), which is mostly done through pointwise numerical calculations of implications.

As it has been mentioned in previous chapters, the basic rules of reasoning used in classical logic are the modus ponens for forward reasoning and modus tollens for backwards reasoning. The symbolic expression of the modus ponens,

$$\begin{array}{lcr} A & \rightarrow & B \\ A & & \\ \hline & & B, \end{array}$$

means that if the rule "if A then B" is given and the event A is observed, then the event B should also be observed.

These processes are referred to in the literature as *inference*. As it has been presented in this book, in the case of fuzzy logic, a generalization of modus ponens is used, based on fuzzy sets. Given a universal set X, a fuzzy set A on X is defined by its membership function $\mu_A : X \rightarrow [0, 1]$ and for all $x \in X$, $\mu_A(x)$ gives the degree of membership of x to A, or the degree with which x fulfills the concept represented by A. Fuzzy sets have a semantic role: they represent the way in which a statement "x is A" is used in a given context.

The generalized modus ponens used in fuzzy control may be given in its simplest expression as

$$\begin{array}{lcr} A & \rightarrow & B \\ A^* & & \\ \hline & & B^*, \end{array}$$

where A, A^*, B and B^* are fuzzy sets, A and A^* are defined on a same universe, but they are not necessarily equal. Similarly for B and B^*. The meaning in this case is the following: given a rule "if A then B" and observing an event A^* which is similar to A, an event B^* is expected, which should also be similar to B.

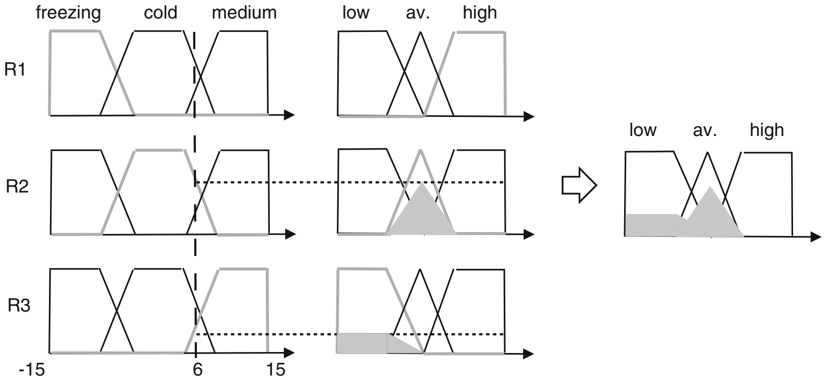
To allow more specified situations to define a rule base in a multivariate system, in the expression $A \rightarrow B$, A may stand for a set of conditions that have to be fulfilled at the same time. The formal representation is a conjunction of fuzzy premises.

For the computation of conjunctions, operations belonging to the class of triangular norms, or simply t-norms, are used. If $T : [0, 1]^2 \rightarrow [0, 1]$ is a t-norm, then it is non-decreasing, associative, commutative and has 1 as identity. The other needed operation is the "then" in "if A then B". In the case of fuzzy logic this operation is not only an extension of the classical implication and its characterization has been thoroughly studied.

Example 8.1.1 The following example is presented at a "phenomenological" level to allow an intuitive understanding of the main issues. In this case a heating system has to be controlled. The system is based on warm water circulating at constant speed and passing through well distributed heating panels. The water temperature should be proportional to the heating demand. The following rules reflect the knowledge of an experienced operator of the heating system:

R1: IF the external temperature is freezing,
THEN the heating demand is large
R2: IF the external temperature is cold,
THEN the heating demand is average
R3: IF the external temperature is medium,
THEN the heating demand is low

It becomes apparent that the meaning of the predicates freezing, cold and medium is different if the system is intended for Ottawa, Madrid or Dakar, since they would not be used in exactly the same way. Similarly, the meaning of a large, average or low heating demand (in a KW scale) is different for a system meant to heat an office room, a conference hall for 200 people or a 25 stores office building. The relative ordering of the terms large, average and low will certainly be the same, and their shapes will probably be the same, in all three mentioned cases. Considering one instance of the problem as illustrated in the figure where it is assumed that the external temperature is 6 °C.



It is fairly obvious that if the external temperature is 6 °C, the first rule does not apply, since 6 ° is not considered to be freezing, but "more cold than medium". The degree of satisfaction of the premises or conditions stated in the "if" part of rules 2 and 3 will affect the strength of the corresponding conclusions. This will be specified by the "then" operation. Rules 2 and 3 are activated to a certain degree and give proportional suggestions for action. These have to be combined into one single action by means of an aggregation operation. From the many aggregation operations that may be used for this purpose, the pointwise maximum is usually the first choice. The aggregated result is also shown in the figure.

Since it was assumed that finally the water temperature should be proportional to the heating demand, the fuzzy set representing the aggregation of activated heating demands has to be converted into a real value through the deffuzzification. This process will imply an information loss, since it is analogous to representing a signal with only one coefficient of its Fourier power spectrum. However, experimental results have shown that "approximating" a fuzzy set by the abscise of its gravity center or the abscise of its center of area leads to an adequate control performance.

8.1.1 Note

This chapter is divided in two main parts that on the one hand pretends to put fuzzy control in the context of the book, and on the other hand introduces the present situation in the area of fuzzy control.

The first part makes a revision of the fundamentals of approximated reasoning explained throughout the book. A condensed view of how to apply the presented theory from a fuzzy control perspective is given.

The second part outlines a glimpse of the state of the art of fuzzy control, where fuzzy plant models and the procedures to obtain them are introduced. In the area of fuzzy control there has been a shift from its original motivation of interpretable fuzzy systems (systems that emulate human control strategy and are easy to use and understand) towards a much more rigorous analysis where mainstream (nonlinear) control criteria are considered. The reader must be aware that in order to fully understand this part, he/she should probably refer to more specialized books in the topic as this section only pretends to present a very general view of the problem.

8.2 Revising Conditional and Implications in Fuzzy Control

The success of fuzzy logic mainly resides in the representation of elementary statements "x is P" ($x \in X$ and P a precise or imprecise predicate or linguistic label on X) by a function $\mu_P : X \to [0, 1]$, in the hypothesis that $\mu_P(x)$ is the degree up to which "x is P", or x verifies the property named P. In the same view, a rule "If x is P, then y is Q" ($x \in X$, $y \in Y$) is represented by a function of the variables μ_P and μ_Q, by

$$R(\mu_P, \mu_Q)(x, y),$$

a number in [0, 1] once P, Q, x and y are fixed.

Functions R can be or can be not functionally expressible, that is, it can exist or it cannot exist a numerical function $J : [0, 1] \times [0, 1] \to [0, 1]$ such that

$$R(\mu_P, \mu_Q)(x, y) = J(\mu_P(x), \mu_Q(y))$$

for all $x \in X$, $y \in Y$. Fuzzy logic works mainly within the positive supposition, and several families of such functions J exist. Such numerical functions are called conditional functions, and their diverse types are derived from the linguistic meaning of the conditional phrase (the rule) "If x is P, then y is Q", that is, from its use in the universe $X \times Y$ at each particular problem.

Even if it is some repetition of what was earlier presented, let's summarize the steps that are necessary.

8.2.1 Inference from Imprecise Rules

If a dynamic systems is considered, with input variables $(x_1, \dots, x_n)$ and output variable y, whose behavior is known by m imprecise rules R_i:

$$
\begin{aligned}
R_1 &: \text{ If } x_1 \text{ is } P_{11} \text{ and } x_2 \text{ is } P_{12} \cdots \text{ and } x_n \text{ is } P_{1n}, \text{ then } y \text{ is } Q_1 \\
R_2 &: \text{ If } x_1 \text{ is } P_{21} \text{ and } x_2 \text{ is } P_{22} \cdots \text{ and } x_n \text{ is } P_{2n}, \text{ then } y \text{ is } Q_2 \\
&\vdots \\
R_m &: \text{ If } x_1 \text{ is } P_{m1} \text{ and } x_2 \text{ is } P_{m2} \cdots \text{ and } x_n \text{ is } P_{mn}, \text{ then } y \text{ is } Q_m
\end{aligned}
$$

The question is the following: If we observe that the input variables are in the "states" "x_1 is P_1^*", "x_2 is P_2^*", ..., "x_n is P_n^*", respectively, what can be inferred for variable y? That is, supposing that it lies in the "state" "y is Q^*", what is Q^*?

Without loss of generality let us consider only the case $n = m = 2$:

$$
\begin{array}{ll}
R_1\text{: If } x_1 \text{ is } P_{11} \text{ and } x_2 \text{ is } P_{12}, & \text{then } y \text{ is } Q_1 \\
R_2\text{: If } x_1 \text{ is } P_{21} \text{ and } x_2 \text{ is } P_{22}, & \text{then } y \text{ is } Q_2 \\
\hline
x_1 \text{ is } P_1^* \text{ and } x_2 \text{ is } P_2^* & \\
\hline
 & y \text{ is } Q^*, \text{ what is } Q^*?
\end{array}
$$

where $x_1 \in X_1$, $x_2 \in X_2$, $y \in Y$, and let us represent it in fuzzy logic by means of the functions R_1^* and R_2^*. This leads to:

$$
\begin{array}{ll}
R_1^*(\mu_{P_{11}} \cdot \mu_{P_{12}}, \mu_{Q_1})((x_1, x_2), y) = & J_1(T_1(\mu_{P_{11}}(x_1), \mu_{P_{12}}(x_2)), \mu_{Q_1}(y)) \\
R_2^*(\mu_{P_{21}} \cdot \mu_{P_{22}}, \mu_{Q_2})((x_1, x_2), y) = & J_2(T_2(\mu_{P_{21}}(x_1), \mu_{P_{22}}(x_2)), \mu_{Q_2}(y)) \\
\hline
\mu_{P_1^*}(x_1) \text{ and } \mu_{P_2^*}(x_2) & \\
\hline
 & \mu_{Q^*}(y) = ?
\end{array}
$$

for convenient continuous t-norms T_1, T_2 and convenient implication functions J_1, J_2. Convenient, in the sense of adequate to the use of the conditional phrases R_1 and R_2 relative to the problem under consideration within the given system S.

To find a solution to this problem, the problem of linguistic control, that is a part of what can be called intelligent systems control, we need to pass throughout several steps.

First Step: Functions *J*

Implication functions $J : [0, 1] \times [0, 1] \to [0, 1]$ are obtained through the interpretation and representation of the rule's use, that is, from the actual meaning of these conditional phrases.Thus, different types of protoform exist

$$a \to b \Leftrightarrow J(a, b)$$

and are used as introduced previously in the book. For example:

- S-implications: $a \to b = a' + b$ and $J(a, b) = S(N(a), b)$ for all a, b in [0, 1], with some strong negation function N and some continuous t-conorm S.
- R-implications: $a \to b = a' + b$ and $J(a, b) = Sup\{z \in [0, 1]; T(z, a) \leqslant b\} = J_T(a, b))$ for all a, b in [0, 1] and some continuous t-norm T.
- Q-implications: $a \to b = a' + a \cdot b$ and J belongs to the family $J(a, b) = S(N(a), T(a, b)$.
- ML-implications: $a \to b = a \cdot b$ and J is in the family of functions $J(a, b) = T(a, b)$.

In fuzzy control mostly ML-implications are considered. Since T is a continuous t-norm, it is $T = Min$, or $T = Prod_\varphi = \varphi^{-1} \circ Prod \circ (\varphi \times \varphi)$, or $T = W_\varphi = \varphi^{-1} \circ W \circ (\varphi \times \varphi)$, with φ an order-automorphism of the unit interval $([0, 1], \leqslant)$, $Prod(x, y) = x \cdot y$ and $W(x, y) = Max(0, x + y - 1)$. Of course, T can be also an ordinal-sum, but these t-norms have never been considered in fuzzy logic. Hence, a ML-implication belongs to the types: $J_M(a, b) = Min(a, b)$, $J_L(a, b) = Prod_\varphi(a, b)$ and $J_W(a, b) = W_\varphi(a, b)$. Only in the third type we can have $J(a, b) = 0$ with $a \neq 0$ and $b \neq 0$, since it is $W_\varphi(a, b) = 0$ whenever $\varphi(a) + \varphi(b) \leqslant 1$, and as in fuzzy control it is desirable not only that $a = 0$ implies $J(a, b) = 0$ but also that $a \neq 0$ and $b \neq 0$ imply $J(a, b) \neq 0$, the third type is rarely used and only $J_M(a, b) = Min(a, b)$ (Mamdani implication) and $J_L(a, b) = a \cdot b$ (Larsen implication) are almost always considered. Notice that for all S, and Q implication functions it is: $J(0, b) = 1$.

Second Step: Modus Ponens

Rules are used in our problem to infer μ_{Q^*}, and this inference requires that when the states of the input variables $x_1, \ldots, x_n$ are exactly those appearing in the antecedent part of one of the m rules, say rule number i, then μ_{Q^*} should be the consequent μ_{Q_i} of this rule. That is, each rule should satisfy the meta-rule of Modus Ponens:

$$
\begin{array}{ll|ll}
\text{If } x \text{ is } P, \text{ then} & y \text{ is } Q & R(\mu_P, \mu_Q) & (x, y) \\
x \text{ is } P^* & & \mu_{P^*}(x) & \\
\hline
 & y \text{ is } Q^* & & \mu_{Q^*}(y)
\end{array}
$$

This meta-rule is satisfied when there is a continuous t-norm T_1 such that

$$T_1(\mu_{P^*}(x), R(\mu_P, \mu_Q)(x, y)) \leqslant \mu_{Q^*}(y),$$

for all $x \in X$, $y \in Y$. When $R(\mu_P, \mu_Q)(x, y) = J(\mu_P(x), \mu_Q(y))$, the last inequation is

$$T_1(\mu_{P^*}(x), J(\mu_P(x), \mu_Q(y)) \leqslant \mu_{Q^*}(y),$$

for all $x \in X$, $y \in Y$.

Hence, for each type of implication function J we need to know which T_1 allows the verification of the Modus Ponens inequality:

$$T_1(a, J(a, b)) \leqslant b, \text{ for all } a, b \in [0, 1].$$

For example, with an S-implication, since $T_1(a, S(N(a), b)) \leqslant b$ implies (with $b = 0$) $T_1(a, N(a)) = 0$, it should be $T_1 = W_\varphi$ for some automorphism φ of $([0, 1], \leqslant)$.

With R-implications J_T, since

$$T(a, J_T(a, b)) = Min(a, b) \leqslant b,$$

the same T in J_T allows to have that inequality.

With Q-implications, since

$$T_1(a, S(N(a), T(a, b))) \leqslant b$$

also implies $T_1(a, N(a)) = 0$ for all $a \in [0, 1]$, it should be also $T_1 = W_\varphi$.

Concerning ML-implications, since

$$T_1(a, T(a, b)) \leqslant Min(a, Min(a; b)) = Min(a, b) \leqslant b,$$

because both $T_1 \leqslant Min$ and $T \leqslant Min$, the Modus Ponens inequality is verified for all t-norms T_1 and, hence, for $T_1 = Min$ (the biggest t-norm).

If T_1 verifies $T_1(a, J(a, b)) \leqslant b$, because of the well-known result that for left-continuous t-norms T_1, $T_1(a, t) \leqslant b$ is equivalent to $t \leqslant J_{T_1}(a, b)$, it results that the inequality is equivalent to $J(a, b) \leqslant J_{T_1}(a, b)$. Hence, among the functions J verifying the Modus Ponens inequality with a continuous t-norm T_1, the R-implication J_{T_1} is the biggest one and, consequently, $T_1(a, J_{T_1}(a, b))$ is closer to b than $T_1(a, J(a, b))$. In particular, it is

$$J_M(a, b) = Min(a, b) \leqslant J_{Min}(a, b) = \begin{cases} 1 & \text{if } a \leqslant b \\ b & \text{if } a > b \end{cases}$$

and

$$J_L(a, b) = a \cdot b \leqslant J_{Min}(a, b) \leqslant J_{Prod}(a, b) = \begin{cases} 1 & \text{if } a \leqslant b \\ \frac{b}{a} & \text{if } a > b \end{cases}$$

(since $a \cdot b \leqslant b$).

Third Step: Zadeh's Compositional Rule of Inference CRI

Once the rule "If x is P, then y is Q" is represented by $J(\mu_P(x), \mu_Q(y))$, and a continuous t-norm T_1 such that $J \leqslant J_{T_1}$ is known, the inference:

$$\begin{array}{l} \text{If } x \text{ is } P, \quad \text{then } y \text{ is } Q \\ \quad x \text{ is } P^* \\ \hline \qquad\qquad\qquad y \text{ is } Q^* \end{array}$$

is obtained by the Zadeh's Compositional Rule of Inference (CRI):

$$\mu_{Q^*}(y) = \operatorname*{Sup}_{x \in X} T_1(\mu_{P^*}(x), J(\mu_P(x), \mu_Q(y))), \text{ for all } y \in Y.$$

It should be pointed out that Zadeh's CRI is not a "result" but a meta-rule. It is a "directive" allowing to reach a solution to our problem, and it should be noticed that when $P^* = P$ it is not in general $Q^* = Q$. For example, in the case of *ML*-implications it is:

$$\mu_{Q^*}(y) = \underset{x \in X}{SupMin}(\mu_P(x), T(\mu_P(x), \mu_Q(y)))$$
$$\leqslant \underset{x \in X}{Sup}\, T(\mu_P(x), \mu_Q(y)) = T(\underset{x \in X}{Sup}\, \mu_P(x), \mu_Q(y)) = \mu_Q(y),$$

provided that $Sup\ \mu_P = 1$, and because of $T(\mu_P(x), \mu_Q(y)) \leqslant \mu_Q(y)$ and T is continuous. But, for example, if $T = Min$, $Sup\ \mu_P = 0.9$ and $Sup\ \mu_Q = 1$, then $\mu_{Q^*}(y) = Min(0.9, \mu_Q(y)) \neq \mu_Q(y)$. Notice that for all the cases in which μ_P is normalized ($\mu_P(x_0) = 1$ for some $x_0 \in X$), Mamdani-Larsen implications do verify $\mu_{Q^*} = \mu_Q$ whenever $\mu_{P^*} = \mu_P$.

Fourth Step: Numerical Input
This is the case in which μ_{P^*} is exactly $x = x_0$ or $x \in \{x_0\}$. That is "x is P^*" is the statement "x is x_0" and hence

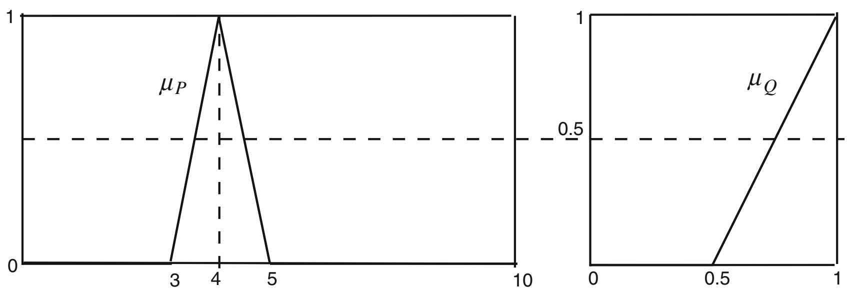
$$\mu_{P^*}(x) = \begin{cases} 1 & \text{if } x = x_0 \\ 0 & \text{if } x \neq x_0 \end{cases}$$

In that case,

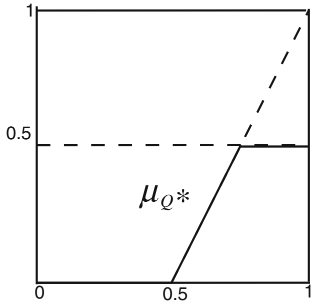
$$\mu_{Q^*}(y) = \underset{x \in X}{Sup} T_1(\mu_{P_j^*}(x), J(\mu_P(x), \mu_Q(y))) = J(\mu_P(x_0), \mu_Q(y)),$$

for all $y \in Y$.

For example, let J be a *ML*-implication, then $\mu_{Q^*}(y) = T(\mu_P(x_0), \mu_Q(y))$. If $X = [0, 10]$, $Y = [0, 1]$, $P = \textit{close to } 4$, $Q = \textit{big}$, with uses as shown in the following figure and moreover $x_0 = 3.5$, with $J(a, b) = Min(a, b)$, then $\mu_{Q^*}(y) = Min(\mu_P(3.5), \mu_Q(y)) = Min(0.5, \mu_Q(y))$, with $\mu_P(x) = x - 3$ between 3 and 4.



Hence, the graphic of the output μ_{Q^*} is the one shown figure



when $x_0 = 3.5$.

Fifth Step: Numerical Consequent
This is the case in which μ_Q is $y = y_0$ or $y \in \{y_0\}$. That is "y is Q" corresponds to "y is y_0", and

$$\mu_Q(y) = \mu_{y_0} = \begin{cases} 1 & \text{if } \; y = y_0 \\ 0 & \text{if } \; y \neq y_0 \end{cases}$$

In this case:

$$J(\mu_P(x), \mu_{y_0}(y)) = \begin{cases} J(\mu_P(x), 1) & \text{if } \; y = y_0 \\ J(\mu_P(x), 0) & \text{if } \; y \neq y_0, \end{cases}$$

and the output μ_{Q^*} depends on the values of J, namely, on $J(a, 1)$ and $J(a, 0)$. Notice that if:

- J is an S-implication, $J(a, 1) = 1$; $J(a, 0) = N(a)$.
- J is an Q-implication, $J(a, 1) = S(N(a), a)$; $J(a, 0) = N(a)$.
- J is an R-implication, $J(a, 1) = 1$; $J(a, 0) = Sup\{z \in [0, 1]; T(z, a) = 0\}$.
- J is an ML-implication, $J(a, 1) = a$; $J(a, 0) = 0$.

When J is an ML-implication, with $J(a, 1) = a$ and $J(a, 0) = 0$:

$$\begin{aligned} \mu_{Q^*}(y) &= \underset{x \in X}{Sup}\, Min(\mu_{P^*}(x), J(\mu_P(x), \mu_{y_0}(y))) \\ &= \mu_{Q^*} = \begin{cases} \underset{x \in X}{Sup}\, Min(\mu_{P^*}(x), \mu_P(x)), & \text{if } \; y = y_0 \\ 0, & \text{if } \; y \neq y_0, \end{cases} \end{aligned}$$

Notice that if the input is also numerical, i.e. $x = x_0$ then the output is:

$$\mu_{Q^*} = \begin{cases} \underset{x \in X}{Sup}\, \mu_{P^*}(x) & \text{if } \; y = y_0 \; \text{ and } \; x = x_0 \\ 0 & \text{if } \; y \neq y_0 \; \text{ and } \; x \neq x_0. \end{cases}$$

Sixth Step: Several Rules with Numerical Input

The problem is now

$$\begin{array}{ll} R_1\colon & \text{If } x \text{ is } P_1, \quad \text{then } y \text{ is } Q_1 \\ R_2\colon & \text{If } x \text{ is } P_2, \quad \text{then } y \text{ is } Q_2 \\ \hline & x \text{ is } P^* \\ \hline & \qquad\qquad\qquad y \text{ is } Q^*? \end{array}$$

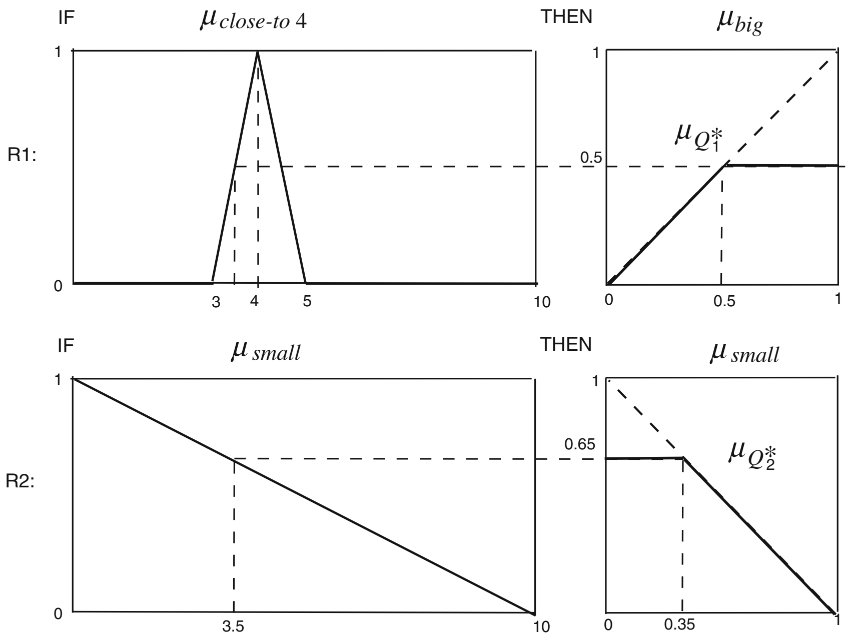
By using the CRI, R_1 with the input "x is P^*" gives an output "y is Q_1^*", with $\mu_{Q_1^*}$ its membership function and R_2 with the input "x is P^*" gives "y is Q_2^*" with $\mu_{Q_2^*}$ its membership function. The total output, "y is Q^*", corresponds to the idea (y is Q_1^*) *or* (y is Q_2^*), and translating this '*or*' by means of the lowest t-conorm its value can be obtained by $\mu_{Q^*}(y) = Max(\mu_{Q_1^*}(y), \mu_{Q_2^*}(y))$.

For example, the Mandani's method consist in taking $J(a, b) = Min(a, b)$ and the Larsen's method, in taking $J(a, b) = a \cdot b$. Let us consider in $X = [0, 10]$, $Y = [0, 1]$ the problem:

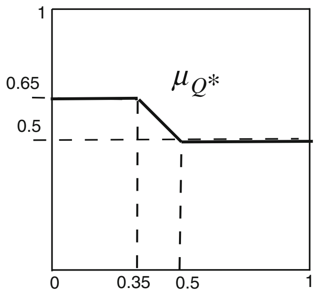
$$\begin{array}{l} R_1 : \text{If } x \text{ is } \textit{close-to } 4, \text{ then } y \text{ is } \textit{big} \\ R_2 : \text{If } x \text{ is } \textit{small}, \qquad \text{then } y \text{ is } \textit{small} \\ \hline \quad x = 3.5 \end{array}$$

and let us find μ_{Q^*} using both methods by supposing "*close-to* 4" as in the former example, $\mu_{big}(y) = y$, $\mu_{small}(x) = 1 - \frac{x}{10}$, and $\mu_{small}(y) = 1 - y$.

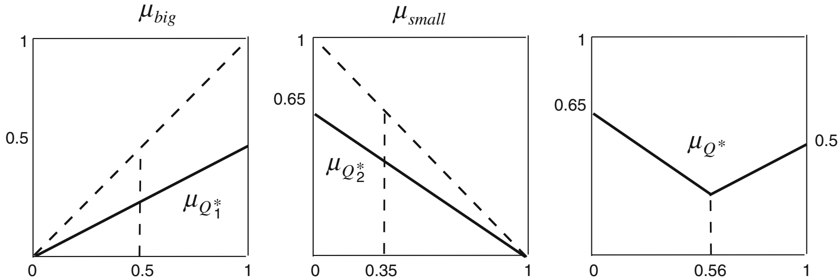
The outputs Q_1^*, Q_2^* are:



- With Mamdani's method: $\mu_{Q_1^*}(y) = Min(\mu_{close\text{-}to\ 4}(3.5), \mu_{big}(y)) = Min(0.5, y)$ (see the figure above, upper right corner)
  $\mu_{Q_2^*}(y) = Min(\mu_{small}(3.5), \mu_{small}(y)) = Min(0.65, 1 - y)$ (see lower right corner of previous figure)
  Hence, $\mu_{Q^*}(y) = Max(\mu_{Q_1^*}(y), \mu_{Q_2^*}(y)) = Max(Min(0.5, y), Min(0.65, 1 - y))$ (as shown in following figure).



- With Larsen's Method: $\mu_{Q_1^*}(y) = 0.5y$ (left of following figure)
  $\mu_{Q_2^*}(y) = 0.65(1 - y)$ (middle of following figure)
  Hence, $\mu_{Q^*}(y) = Max(0.5y, 0.65(1 - y))$ (right of following figure)



Seventh Step: A More Complex Example with Numerical Inputs and Consequences

Let us consider the case:

$$
\begin{array}{rl}
R_1: & \text{If } x_1 \text{ is } P_{11} \text{ and } x_2 \text{ is } P_{12}, \quad \text{then } y = y_1 \\
R_2: & \text{If } x_2 \text{ is } P_{21} \text{ and } x_2 \text{ is } P_{22}, \quad \text{then } y = y_2 \\
\hline
\multicolumn{2}{c}{x_1 = x_1^* \ , \ x_2 = x_2^*} \\
\hline
 & \mu_{Q^*} = ?
\end{array}
$$

With Larsen's method and translating the and in the antecedents also by $T = Prod$.

- Rule R_1 is represented by $J_L(T(\mu_{P_{11}}(x_1), \mu_{P_{12}}(x_2)), \mu_{y_1}(y))$

$$
= \mu_{P_{11}}(x_1) \cdot \mu_{P_{12}}(x_2) \cdot \mu_{y_1}(y) = \begin{cases} \mu_{P_{11}}(x_1) \cdot \mu_{P_{12}}(x_2), & y = y_1 \\ 0, & y \neq y_1 \end{cases}
$$

- Rule R_2 is represented by $J_L(T(\mu_{P_{21}}(x_1), \mu_{P_{22}}(x_2)), \mu_{y_2}(y))$

$$= \mu_{P_{21}}(x_1) \cdot \mu_{P_{22}}(x_2) \cdot \mu_{y_2}(y) = \begin{cases} \mu_{P_{21}}(x_1) \cdot \mu_{P_{22}}(x_2), & y = y_2 \\ 0, & y \neq y_2 \end{cases}$$

Hence, the corresponding outputs under the CRI are:

$$\mu_{Q_1^*}(y) = \begin{cases} \mu_{P_{11}}(x_1^*) \cdot \mu_{P_{12}}(x_2^*), & y = y_1 \\ 0, & y \neq y_1 \end{cases}$$

$$\mu_{Q_2^*}(y) = \begin{cases} \mu_{P_{21}}(x_1^*) \cdot \mu_{P_{22}}(x_2^*), & y = y_2 \\ 0, & y \neq y_2 \end{cases}$$

Consequently:

$$\mu_{Q^*} = Max(\mu_{Q_1^*}(y), \mu_{Q_2^*}(y)) \begin{cases} \mu_{P_{11}}(x_1^*) \cdot \mu_{P_{12}}(x_2^*), & y = y_1 \\ \mu_{P_{21}}(x_1^*) \cdot \mu_{P_{22}}(x_2^*), & y = y_2 \\ 0, & \text{otherwise} \end{cases}$$

Eight Step: Defuzzification

Zadeh's CRI gives an output function μ_{Q^*}, but what it is frequently needed, mainly in control, is an output number as an "order" to be executed by the system. Hence, this step consists in compacting in the best possible way, in a single real number, the information on the system's behavior contained in μ_{Q^*}. That is, the goal is to *defuzzify* μ_{Q^*}.

In the applications the most interesting cases are those in which, respectively, either μ_{Q^*} is a non-null continuous function, or it is a non-null function at only a finite number of points in Y.

In the first case and among the diverse methods that have been suggested in the literature, that known as "center of gravity" of the area below μ_{Q^*} as well as the one known as "center of area" are perhaps the most used ones. In the second case, if for example,

$$\mu_{Q^*}(y) = \begin{cases} \alpha_1, & \text{if } \; y = y_1 \\ \alpha_2, & \text{if } \; y = y_2 \\ \cdots \\ \alpha_n, & \text{if } \; y = y_n \\ 0, & \text{otherwise} \end{cases}$$

the most popular method of defuzzification is that consisting in taking the weighted mean

$$\mu_{Q^*} = \frac{\alpha_1 \cdot y_1 + \alpha_2 \cdot y_2 + \cdots + \alpha_n \cdot y_n}{\alpha_1 + \alpha_1 + \cdots + \alpha_n}$$

The case of m rules with numerical consequents and numerical inputs, with Larsen's implication function and with defuzzification by the weighted mean, is the basis of the so-called Takagi-Sugeno methods of fuzzy inference of orders 1, 2, 3, ... etc.

Example 8.2.1 In the example shown in sixth step where the output μ_{Q^*} is obtained through Mamdani, the area below μ_{Q^*} is easily computed by $0.35 \times 0.15 + \frac{0.15 \times 0.15}{2} + 0.5 \times 1 = 0.564$. Hence, the center of area is a point $y_0 \in (0, 1)$ such that, the areas to the left and to the right of y_0 are equal, i.e.:

$$\begin{aligned}\frac{0.564}{2} = 0.282 &= \int_0^{y_0} \mu_{Q^*}(y)dy \\ &= \int_0^{0.35} 0.65dy + \int_{0.35}^{y_0} (1-y)dy \\ &= 0.228 + y_0 - 0.35 - \int_{0.35}^{y_0} ydy\end{aligned}$$

as the line joining the points (0.35, 0.65) and (0.5, 0.5) is $z = 1 - y$. Hence:

$$y_0 - \int_{0.35}^{y_0} ydy = 0.282 - 0.228 + 0.35 = 0.404$$

$$y_0 - \left[\frac{y^2}{2}\right]_{0.35}^{y_0} = y_0 - \left(\frac{y_0^2}{2} - \frac{0.35^2}{2}\right) = 0.404$$

gives: $y_0^2 - 2y_0 + 0.686 = 0$, with positive root $y_0 = 0.43919$.

Example 8.2.2 In the case of non-null function at two number of points y_1, y_2 (a two rule system) where weighted mean is taken as the output value y_0:

$$y_0 = \frac{\mu_{P_{11}}(x_1^*)\mu_{P_{12}}(x_2^*2)y_1 + \mu_{P_{21}}(x_1^*)\mu_{P_{22}}(x_2^*)y_2}{\mu_{P_{11}}(x_1^*)\mu_{P_{12}}(x_2^*) + \mu_{P_{21}}(x_1^*)\mu_{P_{22}}(x_2^*)}$$

Provided that $X_1 = X_2 = [0, 1]$, $Y = [0, 10]$, $\mu_{P_{11}}(x_1) = x_1$, $\mu_{P_{12}}(x_2) = 1 - x_2$, $y_1 = 6$, $\mu_{P_{21}}(x_1) = 1 - x_1$, $\mu_{P_{22}}(x_2) = x_2$, $y_2 = 4$, $x_1^* = 0.3$ and $x_2^* = 0.7$, the calculation will be:

$$y_0 = \frac{0.3 \times (1 - 0.7) \times 6 + (1 - 0.3) \times 0.7 \times 4}{0.3 \times (1 - 0.7) + (1 - 0.3) \times 0.7} = \frac{0.54 + 1.96}{0.09 + 0.49} = \frac{2.5}{0.58} = 4.31$$

8.2.2 Takagi-Sugeno of Order 1

The method of Takagi-Sugeno of order n is an immediate generalization of the last example with numerical input, numerical consequents, *ML*-implication and defuzzification by a weighted mean. It is the case in which in place of consequents $y = y_i$, y is taken as a polynomial of degree n in the n variables $x_1, x_2, \ldots, x_n$ and with the given coefficients, appearing in the rules' antecedents. For the sake of brevity, and without any loss of generality, let us consider the case of $m = 2$ rules, with $n = 2$ variables:

$$\begin{array}{lllll}
R_1\colon & \text{If} & x_1 \text{ is } P_{11} \text{ and } x_2 \text{ is } P_{12}, & \text{then} & y = \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 \\
R_2\colon & \text{If} & x_1 \text{ is } P_{21} \text{ and } x_2 \text{ is } P_{22}, & \text{then} & y = \beta_1 x_1 + \beta_2 x_2 + \beta_3 \\
\hline
 & & x_1 = x_1^* \text{ , } x_2 = x_2^* & & \\
\hline
 & & & & \mu_{Q*} = ? \text{ ; } y_0 = ?
\end{array}$$

Let us shorten $y = \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3$ by q_1, and $y = \beta_1 x_1 + \beta_2 x_2 + \beta_3$ by q_2. with $q_1^* = \alpha_1 x_1^* + \alpha_2 x_2^* + \alpha_3$ and $q_2^* = \beta_1 x_1^* + \beta_2 x_2^* + \beta_3$. It follows:

$$\mu_{x_1^* x_2^*}(x_1, x_2) = \begin{cases} 1, & \text{if } x_1 = x_1^* \text{ and } x_2 = x_2^* \\ 0, & \text{otherwise} \end{cases}$$

$$\mu_{q_1}(y) = \begin{cases} 1, & \text{if } q_1 \\ 0, & \text{otherwise} \end{cases}$$

$$\mu_{q_2}(y) = \begin{cases} 1, & \text{if } q_2 \\ 0, & \text{otherwise} \end{cases}$$

As rule R_1 is represented by $J_1 = \mu_{P_{11}}(x_1) \cdot \mu_{P_{12}}(x_2) \cdot \mu_{q_1}(y)$ and rule R_2 by $J_2 = \mu_{P_{21}}(x_1) \cdot \mu_{P_{22}}(x_2) \cdot \mu_{q_2}(y)$, it follows:

$$\begin{aligned}
\mu_{Q_1^*}(y) &= \underset{x \in X,\, y \in Y}{Sup}\; Min(\mu_{x_1^* x_2^*}(x_1, x_2), \mu_{P_{11}}(x_1) \cdot \mu_{P_{12}}(x_2) \cdot \mu_{q_1}(y)) \\
&= \mu_{P_{11}}(x_1^*) \mu_{P_{12}}(x_2^*) \mu_{q_1}(y) = \begin{cases} \mu_{P_{11}}(x_1^*) \mu_{P_{12}}(x_2^*), & \text{if } q_1 \\ 0, & \text{otherwise} \end{cases} \\
\mu_{Q_2^*}(y) &= \mu_{P_{21}}(x_1^*) \mu_{P_{22}}(x_2^*) \mu_{q_2}(y) = \begin{cases} \mu_{P_{21}}(x_1^*) \mu_{P_{22}}(x_2^*), & \text{if } q_2 \\ 0, & \text{otherwise} \end{cases}
\end{aligned}$$

Hence,

$$\mu_{Q^*}(y) = Max(\mu_{Q_1^*}(y), \mu_{Q_2^*}(y)) = \begin{cases} \mu_{P_{11}}(x_1^*) \mu_{P_{12}}(x_2^*), & \text{if } q_1 \\ \mu_{P_{21}}(x_1^*) \mu_{P_{22}}(x_2^*), & \text{if } q_2 \\ 0, & \text{otherwise} \end{cases}$$

And finally:

$$y_0 = \frac{\mu_{P_{11}}(x_1^*)\mu_{P_{12}}(x_2^*)q_1^* + \mu_{P_{21}}(x_1^*)\mu_{P_{22}}(x_2^*)q_2^*}{\mu_{P_{11}}(x_1^*)\mu_{P_{12}}(x_2^*) + \mu_{P_{21}}(x_1^*)\mu_{P_{22}}(x_2^*)}$$

8.3 Control of Nonlinear Systems

8.3.1 State-Space Representation

The state-space representation of a dynamical system is used in control engineering in order to completely describe its behavior by first-order differential equations of the involved state variables. Input variables $u(t) = [u_1(t), \ldots, u_p(t)]^T$, output variables $y(t) = [y_1(t), \ldots, y_m(t)]^T$ and state variables $x(t) = [x_1(t), \ldots, x_n(t)]^T$ are all needed to completely define any dynamical system. The state variables of a system constitute the smallest set of variables which completely determine the state of a system and in practice these variables usually correspond with physical magnitudes. Thus, if the state of any system and its inputs are known at any given time, the future behavior of the system will be determined by its state space representation. The state variables are not always measurable and/or observable.

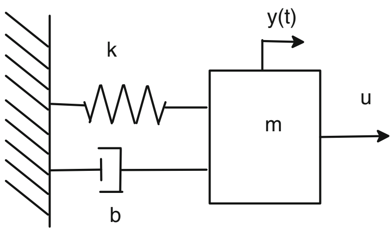
The following equation is a generic state-space representation of a nonlinear system.

$$\begin{cases} \dot{x}(t) = f(x(t)) + g(x(t))u(t) \\ y(t) = h(x(t)) \end{cases}$$

If the dynamical system is linear and time invariant, the differential and algebraic equations may be written in a matrix form, such as,

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t) \\ y(t) = Cx(t) \end{cases}$$

Example 8.3.1 Taken a mechanical system of a mass with a spring and a damper as an example,



where an external force is the only input $u(t)$ to the system, and the position $y(t)$ is the output. The physical equation of this single input single output (SISO) system is

$$m\ddot{y} + b\dot{y} + ky = u$$

which represents a second order linear system. In this case the state variables are defined as

$$x_1(t) = y(t)$$

$$x_2(t) = \dot{y}(t)$$

and so the state space representation is,

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = -\frac{k}{m}x_1 - \frac{b}{m}x_2 + \frac{1}{m}u \end{cases}$$

The output equation is defined by

$$y(t) = x_1(t).$$

In a matrix form, the state equation can be rewritten as,

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{b}{m} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix} u,$$

similarly the output equation is

$$y = \begin{bmatrix} 1 \ 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Both state and output equations form the matrix state space representation of the system

$$\begin{cases} \dot{x} = Ax + Bu \\ y = Cx \end{cases}$$

with

$$A = \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{b}{m} \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix}, \quad C = \begin{bmatrix} 1 \ 0 \end{bmatrix}.$$

8.3.2 Takagi-Sugeno Models for Control of Nonlinear Systems

The so-called Takagi-Sugeno (TS) fuzzy model is particularly useful for the control of nonlinear systems. The TS fuzzy model is described by fuzzy "if-then" rules which represent *local linear* input-output relations of a nonlinear system in a state-space form. The overall fuzzy model of the system is achieved by fuzzy “blending” of the linear systems. A TS fuzzy model representing a continuous dynamical system has the following form,

Rule i:

$$\text{IF } z_1(t) \text{ is } M_{i1} \text{ and } \cdots \text{ and } z_p(t) \text{ is } M_{ip}$$

$$\text{THEN} \begin{cases} \dot{x}(t) = A_i x(t) + B_i u(t) \\ y(t) = C_i x(t) \end{cases}$$

for $i = 1, 2, \ldots, r$.

Here, M_{ij} is the fuzzy set and r is the number of model rules; $x(t) \in R^n$ is the state vector, $u(t) \in R^m$ is the input vector, $y(t) \in R^q$ is the output vector, $A_i \in R^{n\times n}$, $B_i \in R^{n\times m}$, and $C_i \in R^{q\times n}$; $z(t) = \{z_1(t), \ldots, z_p(t)\}$ are known premise variables that may be functions of the state variables and/or external disturbances.

Each linear consequent equation represented by $A_i x(t) + B_i u(t)$ is called a “subsystem”.

$$\begin{cases} \dot{x}(t) = \sum_{i=1}^{r} h_i(z(t))\{A_i x(t) + B_i u(t)\} \\ y(t) = \sum_{i=1}^{r} h_i(z(t))\, C_i x(t) \end{cases}$$

where,

$$h_i(z(t)) = \frac{w_i(z(t))}{\sum_{i=1}^{r} w_i(z(t))},$$

$$w_i(z(t)) = \prod_{j=1}^{p} M_{ij}(z(t))$$

for all t. The term $M_{ij}(z(t))$ is the grade of membership of $z_j(t)$ in M_{ij}. We have $\sum_{i=1}^{r} h_i(z(t)) = 1$, and $h_i(z(t)) \geqslant 0$, for $i = 1, 2, \ldots, r$ and all t.

Thanks to the normalized membership functions, the linear dynamic TS model is in fact a convex combination of local linear models. This property facilitates the stability analysis of the fuzzy system. Fuzzy systems are universal function approximators and hence can be used to model a wide class of processes.

Example 8.3.2 Consider the following nonlinear dynamic system

$$\begin{cases} \dot{x}_1 = -x_1 + x_1 x_2 \\ \dot{x}_2 = x_1 - 3x_2 \\ y \ \ = x_1 \end{cases}$$

with $x_1, x_2 \in [-1, 1]$. This system can be exactly represented, using the sector nonlinearity approach. In this model, the scheduling variable z_1 is chosen as x_2, the fuzzy sets are M_{11} = ‘*around* -1’, M_{12} = ‘*around* 1’, and the corresponding membership functions are $h_{11} = (1 - z_1)/2$ and $h_{12} = (1 + z_1)/2$, respectively. The following TS fuzzy system with linear consequents is the exact representation of the original model,

R_1:

$$\text{IF } z_1 \text{ is } \textit{around } -1$$

$$\text{THEN} \begin{cases} \dot{x} = \begin{pmatrix} -2 & 0 \\ 1 & -3 \end{pmatrix} x \\ y = x_1 \end{cases}$$

R_2:

$$\text{IF } z_1 \text{ is } \textit{around } 1$$

$$\text{THEN} \begin{cases} \dot{x} = \begin{pmatrix} 0 & 0 \\ 1 & -3 \end{pmatrix} x \\ y = x_1 \end{cases}$$

It can be easily seen that this is an exact representation of the nonlinear system in the compact set $x_1, x_2 \in [-1, 1]$, when expressions $\sum_{i=1}^{2} h_i(z(t))\ A_i x(t)$ and $\sum_{i=1}^{2} h_i(z(t))\ C_i x(t)$ are developed,

$$\frac{1-x_2}{2}\begin{pmatrix} -2 & 0 \\ 1 & -3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \frac{1+x_2}{2}\begin{pmatrix} 0 & 0 \\ 1 & -3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -x_1 + x_1x_2 \\ x_1 - 3x_2 \end{pmatrix}$$

$$\frac{1-x_2}{2}x_1 + \frac{1+x_2}{2}x_1 = x_1 = y.$$

The main idea of the controller design based on TS fuzzy systems is to derive each control rule so as to compensate each subsystem or rule of the fuzzy system. So, for every subsystem

$$\dot{x} = A_i x + B_i u$$

a control signal of the form

$$u = -F_i x$$

is selected, resulting in a closed-loop system formed by a set of subsystems such as,

$$\dot{x} = (A_i - B_i F_i)x.$$

The design procedure is conceptually simple and natural. The stability analysis and control design problems can be reduced to linear matrix inequality (LMI) problems.

8.3.3 Stability Analysis

TS fuzzy model is a convex combination of linear models. This structure facilitates stability analysis and observer design by using effective algorithms based on Lyapunov functions and linear matrix inequalities (LMI).

Talking about the Lyapunov function $V(x)$ of a system, is in a way equivalent to talk about the potential energy of the system. For a system to be stable in the sense of Lyapunov, it must remain within a bounded region "close-enough" to an equilibrium point starting from any initial state, and when external input is $u(t) = 0$. Similarly, asymptotic stability means that the system, not only remains close, but must eventually converge to an equilibrium point. Any system in a given initial state must have a positive potential energy $V(x) \geqslant 0$ ($V(x) = 0$ in an equilibrium point), and to ensure that it eventually converges to an equilibrium point, the derivative of this Lyapunov function must be negative $\dot{V}(x) = \frac{d}{dt}V(x) \leqslant 0$. In a sense, it must be proven that the energy of the system eventually dissipates taking the system to an equilibrium point, when no external input is acting over the system ($u(t) = 0$).

The basic stability analysis considers the following quadratic Lyapunov function:

$$V(x) = x^T(t)Px(t),$$

with $P = P^T > 0$. If this Lyapunov function is considered, its derivative along the trajectories of the TS fuzzy subsystem of the type $\dot{x} = A_i x$ is

$$\dot{V}(x) = \dot{x}^T(t)P + P\dot{x}(t) = A_i^T P + PA_i$$

and so:

Theorem 8.3.3 *The equilibrium of a TS fuzzy system is globally asymptotically stable if there exists a common positive definite matrix P such that*

$$A_i^T P + PA_i < 0,$$

for $i = 1, 2, \ldots, r$.

That is, a common P has to exist for all subsystems. This theorem reduces to the Lyapunov stability theorem for continuous linear systems when $r = 1$. If there exists a $P > 0$ such that $V(x(t))$ proves the stability of the system, it is also said to be quadratically stable as $V(x(t))$ is called a quadratic Lyapunov function.

The theorem presents a sufficient condition for the quadratic stability of the TS system, and this is a common P problem that can be solved efficiently via convex optimization techniques for LMIs. For systems and control, the LMI optimization is particularly useful due to the fact that a wide variety of system and control problems can be recast as LMI problems. Apart from a few special cases these problems do not have analytical solutions. However, through the LMI framework they can be efficiently solved numerically in all cases. Therefore, recasting a control problem as an LMI problem is equivalent to finding a "solution" to the original problem.

8.3.4 Parallel Distributed Compensation

The parallel distributed compensation (PDC) offers a procedure to design a fuzzy controller from a given TS fuzzy model. To realize the PDC, a controlled nonlinear system is first represented by a TS fuzzy model. Each control rule is designed from the corresponding rule of a TS fuzzy model, and the designed fuzzy controller shares the same fuzzy sets with the fuzzy model in the premise parts.

The fuzzy controller will have the following form,

Control Rule i:

$$\text{IF } z_1(t) \text{ is } M_{i1} \text{ and } \cdots \text{ and } z_p(t) \text{ is } M_{ip},$$
$$\text{THEN } u(t) = -F_i x(t),$$

for $i = 1, 2, \ldots, r$.

The fuzzy control rules have a linear controller (state feedback laws in this case) in the consequent parts. The overall fuzzy controller is represented by

$$u(t) = -\sum_{i=1}^{r} h_i(z(t)) \; F_i x(t).$$

The fuzzy controller design consists on determining the local feedback gains F_i in the consequent parts. With PDC we have a simple and natural procedure to handle nonlinear control systems. Other nonlinear control techniques require special and rather involved knowledge.

The overall closed-loop system using the PDC controller method is,

$$\dot{x}(t) = \sum_{i=1}^{r}\sum_{j=1}^{r} h_i(z(t))h_j(z(t))\{A_i - B_i F_j\}x(t)\}.$$

For this case, the following sufficient condition for stability can be obtained:

Theorem 8.3.4 *The equilibrium of a fuzzy control system is globally asymptotically stable if there exists a common positive definite matrix P such that*

$$\{A_i - B_i F_j\}^T P + P\{A_i - B_i F_j\} < 0$$

for $h_i(z(t)) \cdot h_j(z(t)) \neq 0$, $\forall t$, $i, j = 1, 2, \ldots, r$.

A control problem in the framework of the TS fuzzy model and PDC design, targets the design of a controller that ensures the stability of the closed-loop system.

Example 8.3.5 Considering the balancing of an inverted pendulum on a cart. The motion equations of the pendulum are described by,

$$\dot{x}_1(t) = x_2(t)$$

$$\dot{x}_2(t) = \frac{g\sin(x_1(t)) - amlx_2^2(t)\sin(2x_1(t))/2 - a\cos(x_1(t))u(t)}{4l/3 - aml\cos^2(x_1(t))}$$

where $x_1(t)$ is the angle of the pendulum, $x_2(t)$ is the angular velocity; g is the gravity, m is the mass of the pendulum, M is the mass of the cart, l is the length of the pendulum, $u(t)$ is the force applied to the cart and $a = 1/(m+M)$.

A fuzzy model which approximates the dynamics of the nonlinear plant for the range $x_1 \in (-\pi/2, \pi/2)$ can be realized by the following two rules,

$$R_1\text{: } \textbf{IF } x_1 \text{ is } \textit{about } 0 \quad \text{THEN } \dot{x}(t) = A_1 x(t) + B_1 u(t)$$

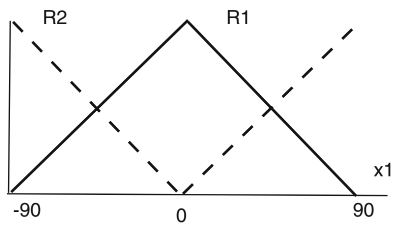
$$R_2\text{: } \textbf{IF } |x_1| \text{ is } \textit{about } \pi/2 \quad \text{THEN } \dot{x}(t) = A_2 x(t) + B_2 u(t)$$

where,

$$A_1 = \begin{bmatrix} 0 & 1 \\ \frac{g}{4l/3 - aml} & 0 \end{bmatrix} \qquad B_1 = \begin{bmatrix} 0 \\ -\frac{g}{4l/3 - aml} \end{bmatrix}$$

$$A_2 = \begin{bmatrix} 0 & 1 \\ \frac{2g}{\pi(4l/3 - aml\beta^2)} & 0 \end{bmatrix} \quad B_2 = \begin{bmatrix} 0 \\ -\frac{a\beta}{4l/3 - aml} \end{bmatrix}$$

and $\beta = \cos(88°)$. The membership functions are



Through the PDC control, the stability of the system is guaranteed for example with

$$F_1 = [-120.67 - 22.67]\,,$$
$$F_2 = [-2551.6 - 764]$$

where it is obtained

$$P = \begin{bmatrix} 3.62 & 0.62 \\ 0.62 & 0.28 \end{bmatrix}.$$

The extension of the pendulum control to the full circle $x_1 \in [-\pi, \pi]$ work space, would be achieved similarly, only by adding two more rules to the TS fuzzy system.

8.3.5 Piecewise Bilinear Model

Recently, Piecewise Bilinear (PB) model is being used for control purpose. The PB model is a fully parametric model to represent Linear/Nonlinear systems. The obtained model is built on piecewise rectangular regions, and each region is defined by four vertices partitioning the state space. As the conventional nonlinear system control based on TS fuzzy model represents a connection of linear state-space models by sector nonlinearity, the PB model represents a convex combination of the vertices defining piecewise regions.

In this approach, bilinear functions are used to regionally approximate any given function. A bilinear function is a nonlinear function of the form $y = a + bx_1 + cx_2 + dx_1x_2$, where any four points in the three dimensional space are spanned with a bi-affine plane.

PB model has a good general approximation capability and it has a continuous crossing over the piecewise regions. Its interpretability, simplicity and visibility facilitates the realization of controllers in industrial applications. A local error does not trigger a global error and its interpolation nature generates robust outputs. A drawback of the PB model is that the stability analysis based on Lyapunov is difficult as bilinear matrix inequalities (BMI) must be solved.

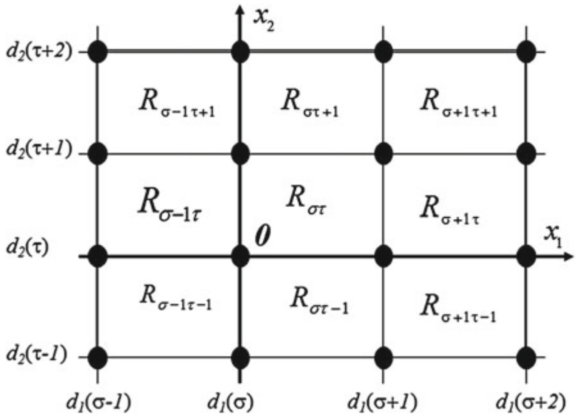
If a general case of an affine two-dimensional nonlinear control system is considered,

$$\begin{cases} \dot{x}_1 = f_1(x_1, x_2) \\ \dot{x}_2 = f_2(x_1, x_2) + g(x_1, x_2) \cdot u \\ y \;\; = h(x_1, x_2) \end{cases}$$

where u is the input. For the PB representation of a state-space equation, a coordinate vector $d(\sigma, \tau)$ of the state space and a rectangle R_{ij} must be defined as,

$$\begin{aligned} d(i, j) &\equiv (d_1(i), d_2(j))^T \\ R_{ij} &\equiv [d_1(i), d_1(i+1)] \times [d_2(j), d_2(j+1)] \end{aligned}$$

where $i \in (1, \ldots, n_1)$ and $j \in (1, \ldots, n_2)$ are integers, and $d_1(i) < d_1(i+1)$, $d_2(j) < d_2(j+1)$. The PB models are formed by matrices of size $(n_1 \times n_2)$, where n_1 and n_2 represent the number of partitions of dimension x_1 and x_2 respectively. Each value in the matrix is referred to as a vertex in the PB model. The operational region of the system is divided into $(n_1 - 1 \times n_2 - 1)$ piecewise regions that are analyzed independently.



The PB model is a particular case of TS systems, as it is derived from a set of fuzzy if-then rules with singleton consequents such that

$$\text{IF } x \text{ is } W^{\sigma\tau}, \quad \text{THEN } \dot{x} \text{ is } f(\sigma,\tau)$$

which in a two-dimensional case, $x \in \mathbb{R}^2$ is a state vector, $W^{\sigma\tau} = (w_1^{\sigma}(x_1), w_2^{\tau}(x_2))^T$ is a membership function vector, $f(\sigma,\tau) = (f_1(\sigma,\tau), f_2(\sigma,\tau))^T \in R$ is a singleton consequent vector, and $\sigma, \tau \in Z$ are integers ($1 \leqslant \sigma \leqslant n_1$, $1 \leqslant \tau \leqslant n_2$) defined by,

$$\sigma(x_1) = d_1(\max(i)) \text{ where } d_1(i) \leqslant x_1,$$

$$\tau(x_2) = d_2(\max(j)) \text{ where } d_2(j) \leqslant x_2.$$

For $x \in R_{\sigma\tau}$, the PB models that approximates a general nonlinear function is expressed as,

$$\begin{cases} f_1(x_1,x_2) = \sum\limits_{i=\sigma}^{\sigma+1}\sum\limits_{j=\tau}^{\tau+1} w_1^i(x_1)w_2^j(x_2)f_1(i,j), \\ f_2(x_1,x_2) = \sum\limits_{i=\sigma}^{\sigma+1}\sum\limits_{j=\tau}^{\tau+1} w_1^i(x_1)w_2^j(x_2)f_2(i,j), \\ g(x_1,x_2) = \sum\limits_{i=\sigma}^{\sigma+1}\sum\limits_{j=\tau}^{\tau+1} w_1^i(x_1)w_2^j(x_2)g(i,j), \\ h(x_1,x_2) = \sum\limits_{i=\sigma}^{\sigma+1}\sum\limits_{j=\tau}^{\tau+1} w_1^i(x_1)w_2^j(x_2)h(i,j), \end{cases}$$

where

$$\begin{cases} w_1^{\sigma}(x_1) &= 1-\alpha, \\ w_1^{\sigma+1}(x_1) &= \alpha, \\ w_2^{\tau}(x_2) &= 1-\beta, \\ w_2^{\tau+1}(x_2) &= \beta, \end{cases}$$

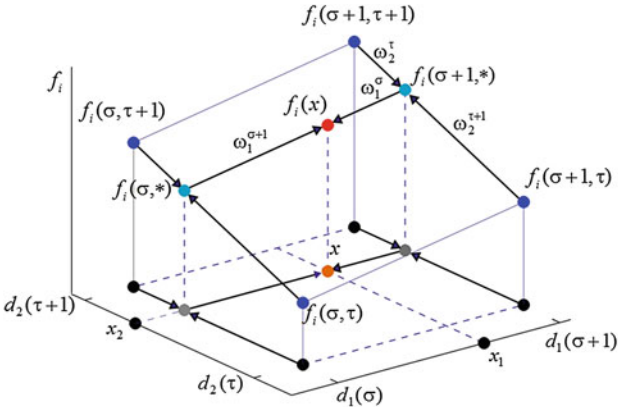
and

$$\alpha = \frac{x_1 - d_1(\sigma)}{d_1(\sigma+1) - d_1(\sigma)}$$

$$\beta = \frac{x_2 - d_2(\tau)}{d_2(\tau+1) - d_2(\tau)}$$

in which case $w_1^i, w_2^j \in [0, 1]$.

In every region of the PB models, i.e.: $f_1(x_1, x_2)$, the values are computed through bilinear interpolation of the corresponding four vertexes as shown in the figure.



Note that the approximation is made by only using the values of a nonlinear function at the vertexes of R_{ij}'s.

8.3.6 Vertex Placement Principle

Generally speaking, if a plant model P and a desired (closed-loop) plant model DP are given, we may have a strategy to design a Controller C by setting a relation $C = DP - P$. This strategy works for linear systems but does not work in general for nonlinear systems as we know. However if nonlinear systems are modeled with PB systems, we can take this strategy by applying Vertex Placement Principle (VPP).

This is a general principle to design a LUT-controller to utilize the characteristics of the PB model of an objective plant. Given a partition of the sate space into sub-regions, the property of a PB model can be completely described by the values of a PB model at the vertexes of regions. Therefore, to design a LUT-controller means to assign the values of a table at the vertex positions. VPP guides us to assign the values of a controller at the vertex positions based on the values of a PB model describing the nonlinear plant. In this sense, VPP is similar to the idea of Pole Placement in linear systems.

Considering an nonlinear plant (P) type

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = f_{2p}(x_1, x_2) + u = \sum_{i=\sigma}^{\sigma+1} \sum_{j=\tau}^{\tau+1} w_1^i(x_1) w_2^j(x_2) f_{2P}(i, j) + u \end{cases}$$

if a controller

$$u = C(x_1, x_2) = \sum_{i=\sigma}^{\sigma+1} \sum_{j=\tau}^{\tau+1} w_1^i(x_1) w_2^j(x_2) C(i, j)$$

is considered, the closed-loop CP will be described as

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = f_{2cp}(x_1, x_2) = f_{2p}(x_1, x_2) + C(x_1, x_2) \end{cases}$$

In order to make the closed-loop behavior of the system be approximated to the performance of a desired plant DP ($f_{2cp} = f_{2dp}$), the controller can be directly computed as,

$$u = C(x_1, x_2) = f_{2dp}(x_1, x_2) - f_{2p}(x_1, x_2).$$

Example 8.3.6 Considering the so called "*Van der Pol*" oscillator, which is a nonlinear plant expressed as,

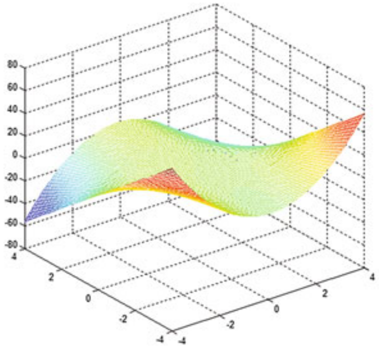
$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = -x_1 + (1 - x_1^2)x_2 + u \end{cases}$$

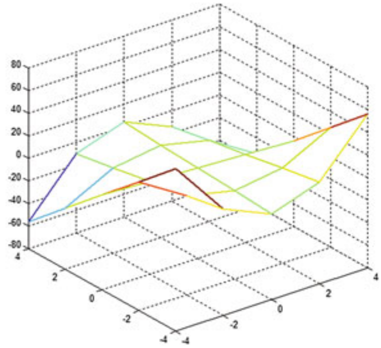
If the operational region of the systems is selected as $-4 \leqslant x_1 \leqslant 4$ and $-4 \leqslant x_2 \leqslant 4$. Thus, the partition of the whole region into piecewise regions is done through $d_1 = d_2 = (-4, -2, 0, 2, 4)$. The piecewise bilinear model representation of the f_{2p} is the following (Table 8.1).

Table 8.1 PB representation of f_{2p}

$x_1 \backslash x_2$	$d_2 = -4$	$d_2 = -2$	$d_2 = 0$	$d_2 = 2$	$d_2 = 4$
$d_1 = -4$	64	34	4	−26	−56
$d_1 = -2$	14	8	2	−4	−10
$d_1 = 0$	−4	−2	0	2	4
$d_1 = 2$	10	4	−2	−8	−14
$d_1 = 4$	56	26	−4	−34	−64

The following figures show both the original surface of the *Van der Pol* system and it's piecewise bilinear representation for the vertex positions d_1 and d_2.





The goal is to design a controller C so that the closed-loop system behaves like the linear plant represented as,

$$\begin{cases} \dot{x_1} = x_2 \\ \dot{x_2} = ax_1 + bx_2 + u \end{cases}$$

where $a = -1$ and $b = -3$. The PB representation of f_{2dp} for the linear plant is the following (Table 8.2).

With the controller calculated as $u = f_{2dp} - f_{2p}$, the following PB model of the controller C is obtained (Table 8.3).

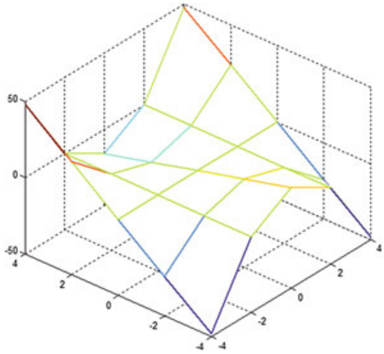
Table 8.2 PB representation of f_{2dp}

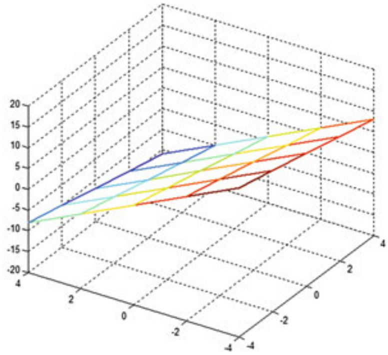
$x_1 \backslash x_2$	$d_2 = -4$	$d_2 = -2$	$d_2 = 0$	$d_2 = 2$	$d_2 = 4$
$d_1 = -4$	16	10	4	−2	−8
$d_1 = -2$	14	8	2	−4	−10
$d_1 = 0$	12	6	0	−6	−12
$d_1 = 2$	10	4	−2	−8	−14
$d_1 = 4$	8	2	−4	−10	−16

Table 8.3 PB model of C

$x_1 \backslash x_2$	$d_2 = -4$	$d_2 = -2$	$d_2 = 0$	$d_2 = 2$	$d_2 = 4$
$d_1 = -4$	–48	–24	0	24	48
$d_1 = -2$	0	0	0	0	0
$d_1 = 0$	16	8	0	–8	–16
$d_1 = 2$	0	0	0	0	0
$d_1 = 4$	–48	–24	0	24	48

Next figures show both the controller C and the resulting closed-loop plant surfaces respectively.





Bibliography

1. B. Bede, *Mathematics of Fuzzy Sets and Fuzzy Logic* (Springer, London, 2013)
2. R. Belohlavek, *Fuzzy Relational Systems: Foundations and Principles* (Kluwer Academic Publishers, New York, 2002)
3. G. Böhme, *Fuzzy-Logik: Einführung in die algebraischen und logischen Grundlagen* (Springer, London, 1993)
4. B. Bouchon-Meunier, *La logique floue* (PUF, Paris, 2003)
5. S.P. Boyd et al., *Linear Matrix Inequalities in System and Control Theory* (SIAM, Philadelphia, 1994)
6. B. Demant, *Fuzzy-theorie oder die faszination des vagen* (Vieweg, Brauschweig, Wiesbaden, 1993)
7. D. Dubois et al., *Fuzzy Sets: History and Basic Notions* (Springer, London, 2000)
8. M. Grabisch et al., *Fuzzy Measures and Integrals: Theory and Applications* (Springer, New York, 2000)
9. G. Klir, B. Yuan, *Fuzzy Sets and Fuzzy Logic*, vol 4 (Prentice Hall, New Jersey, 1995)
10. G.J. Klir, U. St Clair, B. Yuan, *Fuzzy Set Theory: Foundations and Applications* (Prentice-Hall, Upper Saddle River, 1997)
11. Z. Lendek et al., *Stability Analysis and Nonlinear Observer Design Using Takagi-Sugeno Fuzzy Models* (Springer, London, 2010)
12. B. Liu, *Theory and Practice of Uncertain Programming*. (Springer, London, 2009)
13. R. Lowen, *Fuzzy set theory* (Basic Concepts, Techniques and Bibliography, Dordrecht, 1996)
14. R. López de Mántaras, *Approximate Reasoning Models* (Prentice Hall PTR, Upper Saddle River, 1990)
15. D. McNeill, *Fuzzy Logic: The Revolutionary Computer Technology that is Changing Our World* (Simon and Schuster, New York, 1994)
16. G. Metcalfe, N. Olivetti, D.M. Gabbay, *Proof Theory for Fuzzy Logics*, vol 36 (Springer, London, 2008)
17. H.T. Nguyen, E.A. Walker, *A First Course in Fuzzy Logic* (CRC Press, Boca Raton, 2005)
18. A. Syropoulos, *Theory of Fuzzy Computation* (Springer, New York, 2014)
19. K. Tanaka, *An Introduction to Fuzzy Logic for Practical Applications* (Springer, New York, 1997)
20. K. Tanaka, H.O. Wang, *Fuzzy Control Systems Design and Analysis: A Linear Matrix Inequality Approach* (Wiley, Danvers, 2004)
21. E. Trillas, *Conjuntos Borrosos* (Vicens-Vives, Barcelona, 1980)
22. E. Trillas, C. Alsina, J.-M. Terricabras, *Introducción a la lógica borrosa*. (Ariel Barcelona, Spain, 1995)

E. Trillas and L. Eciolaza, *Fuzzy Logic*, Studies in Fuzziness
and Soft Computing 320, DOI 10.1007/978-3-319-14203-6

23. E. Turunen, E. Turunen, *Mathematics Behind Fuzzy Logic* (Physica-Verlag, Heidelberg, 1999)
24. X. Wang, D. Ruan, E.E. Kerre, *Mathematics of Fuzziness: Basic Issues* (Springer, Heidelberg, 2009)
25. J. Yen, R. Langari, *Fuzzy Logic: Intelligence, Control, and Information* (Prentice-Hall, Upper Saddle River, 1998)
26. L.A. Zadeh, *Fuzzy Sets and Applications: Selected Papers* (Wiley, New York, 1987)
27. L.A. Zadeh, *Computing with Words: Principal Concepts and Ideas* (Springer, Berlin, 2012)
28. L.A. Zadeh, B. Yuan, G.J. Klir, *Fuzzy sets, fuzzy logic, and fuzzy systems: selected papers by Lotfi A* (World Scientific, Zadeh, 1996)
29. H.J. Zimmermann, *Fuzzy Set Theory and Its Applications*, Second Revised Edition. (Springer, New York, 1992)