DOTTORATO DI RICERCA

in

SCIENZE COMPUTAZIONALI E INFORMATICHE

Ciclo XVIII

Consorzio tra Università di Catania, Università di Napoli Federico II,
Seconda Università di Napoli, Università di Palermo, Università di Salerno

SEDE AMMINISTRATIVA: UNIVERSITÀ DI NAPOLI FEDERICO II

CRISTINA COPPOLA

DISTANCES AND CLOSENESS MEASURES IN INFORMATION SPACES

TESI DI DOTTORATO DI RICERCA

Table of Contents

Introduction

The notion of metric space plays a basic role in several researches addressed to process information. Indeed the objects we will investigate are represented by points and the distance is a measure of “dissimilarity” between objects. Now, the question that arises is if such a notion is the better one in a context in which we are not able to obtain complete information about the considered objects.

This thesis is devoted to face this question, by giving suitable axioms extending the usual ones for metric spaces and by considering *regions* in a suitable space, instead of the points. This idea originates from A. N. Whitehead’s researches, aimed to define a geometry without the concept of point as primitive (see [46], [47] and [48]) and from a metrical version of these researches, proposed by G. Gerla (see, for example [23], [24]). Indeed, we can re-interpret the regions as ”incomplete pieces of information” and the diameter of a region as a measure of the vagueness of the available information: the bigger it is, the more there is uncertainty. Points (having zero-diameter) represent complete information.

Another idea examined in this thesis is the possibility of referring to the “logical” notion of closeness instead of the one of distance. Indeed, there is a duality between these concepts, that is easily understandable: when comparing objects accordingly to their properties, we can use both a measure of how they are “similar” and a measure of how they are “dissimilar”; the smaller the distance is, the bigger the closeness is. We investigate the notion of closeness in the fuzzy domain, examining similarities and fuzzy orders.

More precisely, the thesis is structured as follows.

In Chapter 1 we first give some necessary basic notions in multi-valued logics. Then we give some information about the metric structures we will start from and we show some already known dualities between the metric notions and the fuzzy relations.

In Chapter 2 we propose an approach to establish a link between point-free geometry and the categorical approach to fuzzy sets theory (as proposed by Höhle in [28]). In particular, starting from the definition of pointless metric spaces, we introduce the *pointless ultrametric spaces*. Then we define the semimetric spaces, the semisimilarities on some spaces, and we verify the relations between these two kind of structures. Moreover, we focus on *semiultrametric spaces* and on the semisimilarity with the *Gödel* t-norm, called *G-semisimilarity*. We also examine the relations existing between pointless metric spaces and semimetric spaces, and, in particular, between pointless ultrametric and pointless semiultrametric spaces. Besides, we verify the connection between the structures equipped with G-semisimilarities and the pointless ultrametric spaces. Finally, once we have organized the class of pointless ultrametric spaces into a category, we define two functors to relate such a category with Höhle's category, and we exhibit a class of examples of G-semisimilarity.

In Chapter 3 we introduce the concept of *approximate distance* in agreement with the ideas of Interval Analysis, (see, for example, [3] and [27]). Approximate distances extend the notion of distance by taking into account errors arising from the incomplete knowledge of the points. We do this by using interval-valued maps (see [6], [7]). Besides, developing Whitehead's ideas, we introduce the approximate distance between regions (see [5]). Hence, considering an interval-valued "distance" Δ on an ordered space of regions, we define an abstract structure of *interval semimetric space*. We interpret a region x as representing the incompleteness of the knowledge and $\Delta(x, y)$ as an approximate measure of how two pieces of information x and y are close; we also define a weight function p, intending $p(x)$ as a measure of the completeness of x. Canonical models of the resulting theory are obtained from classes of bounded subsets of pseudometric spaces by the *minimum* and *maximum* distances. Also we apply the notion of approximate distance to some topics in Computer Science. In particular, we refer to Fuzzy Sets Theory, where two definitions of interval-distances between fuzzy subsets are proposed (namely, by cuts and by hypographs). Then we define interval-distances between rough sets. Finally, as an application to the clustering problem, an algorithm based on interval distances between clusters is examined.

In Chapter 4, among the distances that verifies weaker axioms with respect to the metrics, we take into account the quasi-metrics, in which is not required the symmetry. The dual notion of $*$-fuzzy preorder allows us to extend simultaneously both the notions, metric in nature, and the ones of ordered set theory. In particular, we give fixed points theorems in sets equipped with $*$-fuzzy preorders extending both the theorems in metric spaces and the theorems in ordered sets. Finally some applications to logic programming are suggested.

Acknowledgments

I would like to thank all the people who supported me during my studies. First and foremost, I wish to express my gratitude to my advisor, Professor Giangiacomo Gerla for his invaluable guide during my researches.

I would like to thank Professor Amelia Nobile for her kindness in giving me suggestions. I want also express my gratitude to Dott. Brunella Gerla and Dott. Annamaria Miranda for their guidance during these years.

I wish also to thank Tiziana Pacelli, who shared with me part of the efforts needed for the realization of this work and not only.

I thank my friends and colleagues of the DMI of Salerno for their friendship and their patience also in hard times.

Finally, I thank the little Matteo and Cristina for their sweetness, my guardian angels, all the friends of mine and the people that with their support helped me during my PhD course.

Chapter 1

Basic notions

In this Chapter we give some notions which underlie what we are going to say in the remainder of the dissertation. In particular, in order to provide some tools useful to deal with information and incomplete information, we give some definitions in a logic setting on a side and in a metric one, on the other side.

We start by introducing some algebraic notions on which many–valued logics are based (for a wider study, see for example [26], [35]). Then we examine some distances and we show some dualities between the notions introduced in the two settings.

1.1 Algebraic structures for multi-valued logic

Classical logic is based on the *bivalence*-hypothesis: every proposition is either true or else false whereas "non-determinate" truth values are not taken under consideration. Another of its basic properties is the *truth-functionality* of the logical connectives: the truth value of a compound formula depends on the truth values of its compounds, unambiguously. At the beginning of the XX century some attempts were made in order to formalize many-valued logics, whose truth degrees are not two yet, but three or more. The work of Jan Lukasiewicz (1920) and that one of Heyting (1930) represent some of the first important examples of non-classical logic and, also nowadays, many researchers focus their studies on this kind of topic.

In this section we present some notions concerning the algebraic structures utilized for the evaluation of formulas, both in classical logic and in multi-valued one.

Definition 1.1.1 A *lattice* is a structure $(L, \vee, \wedge, 0, 1)$, where $\vee$ and $\wedge$ are binary operations, satisfying, for every $x, y, z \in L$, the following axioms:

- $x\vee y = y\vee x$; $x\wedge y = y\wedge x$ (commutativity)
- $x\vee(y\vee z) = (x\vee y)\vee z$; $x\wedge (y\wedge z) = (x\wedge y) \wedge z$ (associativity)
- $x\vee x=x$; $x\wedge x= x$ (idempotency)
- $0\vee x= x$; $1\wedge x= x$ (neutral elements)

Lattices coincide with particular kind of ordered sets.

Proposition 1.1.1 *Let* $(L, \vee, \wedge, 0, 1)$ *be a lattice and let* $\leq$ *be an order relation defined by* $x\leq y$ *iff* $x\wedge y= x$. *Then* $(L, \leq, 0, 1)$ *is an ordered set such that* $Inf\{x, y\}= x\wedge y$ *and* $Sup\{x, y\}= x\vee y$ *and* 0 *and 1 are the smallest and the greatest element, respectively. Conversely, let* $(L, \leq, 0, 1)$ *be an order set with a minimum element* 0 *and a maximum element* 1 *and such that, for every* $x, y \in L$, *there exists the supremum* $x\vee y$ *and the infimum* $x\wedge y$. *Then* $(L, \vee, \wedge, 0, 1)$ *is a lattice such that* $x\leq y$ *iff* $x\wedge y= x$.

The order $\leq$ defined for each lattice $(L, \vee, \wedge, 0, 1)$, as in the last proposition, is also called the *order determined by L.* A lattice $(L, \vee, \wedge, 0, 1)$ is *complete* if each $X \subseteq L$ has its *sup* and *inf*. As an example, the real interval [0, 1] with the operations of maximum and minimum is a complete lattice.

Important instances of lattices are distributive lattices and lattices with complements.

Definition 1.1.2 A lattice is called *distributive* if the following identities hold:

$$x\vee(y\wedge z) = (x\vee y)\wedge(x\vee z)$$
$$x\wedge(y\vee z) = (x\wedge y)\vee(x\wedge z).$$

Definition 1.1.3 A *lattice with complements* is a lattice with a unary operation of complement $-$, such that

- $x\vee -x= 1$ (law of the excluded middle)
- $x\wedge -x= 0$ (law of non-contradiction)

Definition 1.1.4 A *Boolean algebra* is a distributive lattice with complements and it is denoted by $(L, \vee, \wedge, -, 0, 1)$.

Two typical examples of Boolean algebra are the *Boolean algebra for classical logic* and the *Boolean algebra of subsets*. The first has the set $\{0,1\}$ as support, with only two elements indicating the two values "false" and "true". The two binary operations are defined by $x\vee y=max\{x,y\}$ and $x\wedge y=min\{x,y\}$ and the unary operation is defined by $-x = 1-x$. The second example is a structure $(P(S), \cup, \cap, -, \emptyset, S)$, where $P(S)$ is the powerset of S and the operations $\cup, \cap, -$ are the usual union, intersection and complement of sets.

In order to evaluate formulas in multi-valued logic, many structures generalizing the Boolean algebra for classical logic, can be considered. An important class is the class of *residuated lattices.*

Definition 1.1.5 A *residuated lattice* is a structure $(L, \vee, \wedge, *, \rightarrow, 0, 1)$ such that

- $(L, \vee, \wedge, 0, 1)$ is a lattice;
- $(L, *, 1)$ is a commutative monoid;
- $*$ is isotone in both arguments, i.e.

$$x \leq y \;\Rightarrow\; x * z \leq y * z,$$

$$x \le y \;\Rightarrow\; z * x \le z * y\,;$$

- $\to$ is a residuation operation with respect to $*$, i.e.

$$x*y \le z \quad \text{iff} \quad x \le y \to z.$$

We say that $(L, \vee, \wedge, *, \to, 0, 1)$ is *complete* provided that the lattice is complete. The operation $*$ is called *multiplication* and $\to$ is called *residuation*. In the case that $L= \{0,1\}$, these operations coincide with the usual minimum and classical implication, respectively.

We are interested to the residuated lattices in which L coincides with [0,1]. In the following, we provide some of the most important examples of such a class, depending on the choice of the operation $*$ and its related residuation.

1.The *Gödel algebra* is the structure $([0,1], \vee, \wedge, \to_G, 0, 1)$, where $* = \wedge$ and

$$x \to_G y = \begin{cases} 1 & \textit{iff} \quad x \le y \\ y & \textit{iff} \quad y < x \end{cases}$$

2. The *Goguen algebra* is the structure $([0,1], \vee, \wedge, *, \to_P, 0, 1)$, where $*$ is the usual product of reals and the residuation is

$$x \to_P y = \begin{cases} 1 & \textit{iff} \quad x \le y \\ \dfrac{y}{x} & \textit{iff} \quad y < x \end{cases}$$

3. The *Lukasiewicz algebra* is the structure $([0,1], \vee, \wedge, *, \to_L, 0, 1)$, where

$$x*y = 0 \vee (x+y-1) \qquad \text{(\textit{Lukasiewicz conjunction})}$$

and

$$x \to_L y = 1 \wedge (1-x+y). \qquad \text{(\textit{Lukasiewicz implication})}$$

The following proposition lists the main properties of a complete residuated lattice (see [35]).

Proposition 1.1.2 *Let* $(L,\vee, \wedge, *, \to, 0, 1)$ *be a complete residuated lattice, let* $x, y, z \in L$ *and* $(x_i)_{i\in I}$ *be a family of elements in L. Then the followings hold true:*

(i) $x\to x = 1$,

(ii) $(x\to y) * (y\to z)\leq x\to z$,

(iii) $x\to y = 1$ and $y\to x = 1 \Rightarrow x = y$,

(iv) $x\to y = 1 \Leftrightarrow x\leq y$,

(v) $x\to y = Sup\{z\in L : x * z \leq y\}$,

(vi) $(z\to y) * z \leq y$,

(vii) $Sup_{i\in I} (x* x_i) = x*(Sup_{i\in I}\, x_i)$,

$(viii)$ $Sup_{i\in I} (x\to x_i) \leq x\to (Sup_{i\in I}\, x_i)$,

(ix) $Sup_{i\in I} (x_i\to x) \leq (Inf_{i\in I}\, x_i)\to x$,

(x) $Inf_{i\in I} (x * x_i) \geq x* (Inf_{i\in I}\, x_i)$,

(xi) $Inf_{i\in I} (x\to x_i) = x\to (Inf_{i\in I}\, x_i)$,

(xii) $Inf_{i\in I} (x_i\to x) = (Sup_{i\in I}\, x_i)\to x$..

1.2 Triangular norms

In the following we present an overview of a particular kind of operations on the real interval [0,1]. In fact they were introduced by Menger (see [32]) and then elaborated by Schweizer and Sklar (see [41]) in order to generalize the concept of the triangular inequality.

Definition 1.2.1 A *triangular norm* (briefly *t-norm*) is a binary operation $*$ on [0,1] such that, for all $x, y, z, x_1, x_2, y_1, y_2 \in [0,1]$

- $*$ is commutative, i.e.,

$$x* y = y * x,$$

- $*$ is associative, i.e.,

$$(x * y) * z = x * (y * z),$$

- $*$ is isotone in both arguments, i.e.,

$$x_1\leq x_2 \Rightarrow x_1* y\leq x_2 * y,$$

$$y_1\leq y_2 \Rightarrow x * y_1\leq x * y_2,$$

- $*$ verifies the *boundary conditions*, i.e.

$$1*x = x = x*1 \text{ and } 0 *x = 0 = x*0.$$

The notion of t-norm is suitable to represent the truth function of the conjunction (see [26]). In fact, the intuitive understanding of the conjunction is the following: given two formulas α and β, a large truth degree of their conjunction $\alpha \wedge \beta$ should indicate that both the truth degree of α and the truth degree of β is large, without any preference between α and β. Moreover, a truth function of connectives has to behave classically for the values 0 and 1, since any multi-valued logic has to be a generalization of the classical logic. The t-norms' properties of isotony in both arguments, of 1 as unit element and 0 as zero element satisfy these requirements.

A t-norm $*$ is *continuous* if it is a continuous map $* : [0, 1]^2 \rightarrow [0,1]$ in the usual sense. The most important examples of continuous t-norms are:

- *Gödel t-norm*: $x * y = min\,(x, y)$, (1.1)
- *Product t-norm* : $x * y = x \cdot y$, in the sense of product of reals, (1.2)
- *Lukasiewicz t-norm*: $x * y = max\,(0, x + y - 1)$. (1.3)

They are fundamental in the sense that each continuous t-norm can be expressed as a combination of them (see [26]).

For each continuous t-norm $*$, representing the truth function of conjunction, it is possible to consider a map, associate to $*$, suitable to represent the truth function of implication.

Definition 1.2.2 Let $*$ be a continuous t–norm. The associated *residuation* is the operation $\rightarrow_*$ defined by

$$x \rightarrow_* y = sup\{a \,/\, x * a \leq y\}.$$

As an immediate consequence, we can observe (see [26]) that, given a continuous t- norm $*$, the residuation $\rightarrow_*$ is the unique operation satisfying the condition

$$x * a \leq y \;\Leftrightarrow\; a \leq x \rightarrow_* y.$$

In classical logic, the implication $\alpha \rightarrow \beta$ is true if and only if the truth-value of α is less than or equal to the truth-value of β. Generalizing, (see [26]) we can say that, in a multi-valued logic, a large truth-value of $\alpha \rightarrow \beta$ should indicate that the truth-value of α is "not too much larger" than the truth value of β. In accordance with this interpretation, the notion of residuation, associated to a t-norm, is adequate to represent the truth function of the implication. In particular, $x \rightarrow_* y=1$ if and only if $x \leq y$ and $1 \rightarrow_* x = x$, as it happens in the two-valued logic.

Let us observe that the three continuous t-norms (1.1), (1.2), (1.3) are just the same operations we have considered in the three algebraic structures defined in Section 1.1. Moreover, if we define the related residuations as in Definition 1.2.2, we obtain the implications of the Gödel, Goguen and Lukasiewicz algebras, respectively.

Definition 1.2.3 Let $*$ be a t-norm. The associated *biresiduation* $\leftrightarrow_*$ is the operation $\leftrightarrow_*: [0,1]^2 \rightarrow [0,1]$ defined by

$$x \leftrightarrow_* y = (x \rightarrow_* y) \wedge (y \rightarrow_* x).$$

The biresiduation operations associated to the three basic norms are so defined, for every $x, y \in [0,1]$:

- for the Gödel t-norm :

$$x \leftrightarrow_G y = \begin{cases} 1 & if\ a = b \\ x \wedge y & otherwise \end{cases}$$

- for the product:

$$x \leftrightarrow_P y = \frac{x \wedge y}{x \vee y} = exp\,(-|\,log\, x - log y\,|),$$

where $0/0=1$ and $\infty - \infty = 0$.

- for the Lukasiewicz t-norm :

$$x \leftrightarrow_P y = 1 - |x - y|.$$

Proposition 1.2.1 *Let* $(L, \vee, \wedge, *, \rightarrow, 0, 1)$ *be a complete residuated lattice, x, y and z be elements in L and let* $\leftrightarrow_*$ *be the biresiduation associated to* $*$. *Let* $(x_i)_{i \in I}$ *a family of elements in L. Then the following hold true:*

(i) $x \leftrightarrow_* x = 1$,

(ii) $x \leftrightarrow_* y = 1 \Leftrightarrow x = y$

(iii) $(x \leftrightarrow_* y) * (y \leftrightarrow_* z) \leq x \leftrightarrow_* z$

(iv) $x \leftrightarrow_* y = y \leftrightarrow_* x$.

Let us consider now a particular class of continuous t-norms, the *Archimedean t-norms*.

Definition 1.2.4 A continuous t–norm $*$ is called *Archimedean* if, for any $x, y \in [0,1]$, $y \neq 0$, an integer n exists such that $x^{(n)} < y$, where $x^{(n)}$ is defined by $x^{(1)} = x$ and $x^{(n+1)} = x^{(n)} * x$.

The usual product and the Lukasiewicz t-norm are examples of Archimedean continuous t-norms, while the minimum is an example of continuous t-norm which is not Archimedean. In order to characterize the Archimedean triangular norms, let us consider the extended interval $[0, \infty]$ and let us assume that $x + \infty = \infty + x = \infty$ and that $x \leq \infty$ for any $x \in [0, \infty]$.

Let f: $[0, 1] \rightarrow [0, \infty]$ be a continuous, strictly decreasing function such that $f(1) = 0$. The function $f^{[-1]} : [0, \infty] \rightarrow [0,1]$ defined by

$$f^{[-1]}(y) = \begin{cases} f^{-1}(y) & \textit{if } \; y \in f([0,1]), \\ 0 & \textit{otherwise} \end{cases}$$

is called *pseudoinverse* of f.

Trivially, $f^{[-1]}$ is order-reversing, $f^{[-1]}(0) = 1$ and $f^{[-1]}(\infty) = 0$. Moreover, for any $x \in [0, 1]$, $f^{[-1]}(f(x)) = x$ and

$$f(f^{[-1]}(x)) = \begin{cases} x & \text{if } x \in f([0,1]), \\ f(0) & \text{otherwise} \end{cases}$$

Definition 1.2.5 Let f: $[0, 1] \to [0, \infty]$ be a continuous, strictly decreasing function such that $f(1) = 0$. Then f is called *additive generator* of a *t*-norm $*$ if it results, for every $x, y \in [0, 1]$,

$$x * y = f^{[-1]}(f(x) + f(y)), \tag{1.4}$$

Proposition 1.2.2 *A function* $*$: $[0,1]^2 \to [0,1]$ *is a continuous Archimedean t-norm iff it has an additive generator.*

As an example, the additive generator of the product t-norm is $f_p(x) = -ln(x)$ and the additive generator of the Lukasiewicz t-norm is $f_L(x) = 1-x$.

Let us observe that, if an additive generator exists for a t-norm $*$, then we can write

$$x \to_* y = f^{[-1]}(f(y) - f(x)).$$

and

$$x \leftrightarrow_* y = f^{[-1]}(|f(x) - f(y)|).$$

1.3 *L*-subsets

An extension of classical logic is fuzzy logic. Classical logic holds that everything can be expressed in binary terms (0 or 1, black or white, yes or no), whereas fuzzy logic replaces boolean truth values with *degrees of truth*. In fuzzy logic the basic notion, which allows us to describe these degrees of truth, is that of *fuzzy set* (see [52]). Fuzzy set theory can be regarded as an extension of the classical set theory, where the membership of elements to a set is estimate in binary terms: an element either belongs or does not belong to the set. Fuzzy set theory permits the gradual assessment of the membership of elements to a set. This property is described by a membership function.

Let X be a classical set of objects, called the *universe*, and the let x be a generic element of X. A subset A of X can be represented by its *characteristic function* from X to $\{0,1\}$

$$\chi_A(x) = \begin{cases} 1 & iff \quad x \in A \\ 0 & iff \quad x \notin A \end{cases}$$

In order to construct a generalized characteristic function, we consider any complete lattice L.

Definition 1.3.1 Let L be a complete lattice and let S be a set. We call *L-subset* of S any map s: $S \to L$ and we denote by L^S or by $\Im$ (S) the class of all the L-subsets of S. If L is the lattice [0,1] the map s is called *fuzzy subset* of S.

Given any x in S, the value $s(x)$ is the "degree of membership" of x to s. In particular, $s(x)=0$ means that x is not included in s, whereas 1 is assigned to the elements fully belonging to s. The values between 0 and 1 characterize the elements with a "non well-defined" membership: the closer to 1 the value of $s(x)$ is, the more x belongs to s. Any L-subset s such that $s(x) \in \{0, 1\}$, for any $x \in S$, is called *crisp set*. The *support* of a fuzzy set s is the classical subset of X, *supp* s = $\{x \in S / s(x) > 0\}$. The *height* of s is defined by $hgt(s) = sup\{s(x) / x \in S\}$. We say that a fuzzy set is *normal* if there exists an element x of S such that $s(x)=1$.

Now we examine how the basic notions of set theory can be naturally extended to the fuzzy subsets.

Definition 1.3.2 We define the *inclusion* relation $\subseteq$ by setting, for any s, $s' \in \Im(S)$ and for every $x \in S$,

$$s \subseteq s' \Leftrightarrow s(x) \leq s'(x).$$

If $s \subseteq s'$ we say that s is *contained* in s' or that s is a *part* of s'.

Definition 1.3.3 We define the *union* $s \cup s'$, the *intersection* $s \cap s'$ and the *complement* $\sim s$ by setting, for any s, $s' \in \Im(S)$ and for every $x \in S$,

$$(s \cup s')(x) = s(x) \vee s'(x)$$

$$(s \cap s')(x) = s(x) \wedge s'(x).$$

$$(\sim s)(x) = -s(x).$$

Let us observe that, in accordance with Zadeh's formulas, proposed in 1965, the symbols $\vee$, $\wedge$, $-$ are the usual supremum, infimum, complement operations in a lattice L, respectively. In such a case, when the lattice is restricted to the set $\{0,1\}$, the formulas of the last definition give us the usual union, intersection and complement of ordinary sets. Let us stress that there are many other ways of extending these basic operations and a very reasonable proposal for defining intersection between fuzzy sets is given by t-norms, introduced in Section 1.2.

Proposition 1.3.1 *The structure* $(\Im(S), \cup, \cap, \sim, s^0, s^1)$ *is a complete, completely distributive lattice which extends the Boolean algebra* $(P(S), \cup, \cap, \sim, \varnothing, S)$.

Proof. Indeed we can associate any subset X of S with the related characteristic function χ_X. More precisely, the map H: $P(S) \rightarrow \Im(S)$, defined by setting $H(X) = \chi_X$ for any $X \in P(S)$, is an injective lattice homomorphism from $P(S)$ to $\Im(S)$. So we identify the classical subsets of S with the crisp L-subsets of S. Particularly, we identify $\varnothing$ with the map s^0 constantly equal to 0 and S with the map s^1 constantly equal to 1. In general, given $\lambda \in L$, we indicate by s^λ the map constantly equal to λ.

□

For exhibiting an element x of S that typically belongs to a fuzzy set s, we can require its membership value $s(x)$ to be greater than some threshold $\lambda \in L$. The ordinary set of such elements is the *closed* λ *- cut* of s

$$C(s, \lambda) = \{x \in S \,/\, s(x) \geq \lambda\}.$$

Analogously, we can define the *open* λ*-cut* of an L-subset s of S setting $O(s, \lambda) = \{x \in S \,/\, s(x) > \lambda\}$. Let us observe that an L-subset s of S can be expressed in terms of the characteristic functions of its λ-cuts

$$s(x) = sup\ \{min\ (\lambda\ , \chi_{C(s,\lambda)}(x)) \,/\, \lambda \in L\},$$

where $\chi_{C(s,\lambda)}(x) = \begin{cases} 1 & iff \quad x \in C(s,\lambda) \\ 0 & iff \quad x \notin C(s,\lambda) \end{cases}$.

The following properties about λ-cuts can be easily deduced.

Proposition 1.3.2 *Let s, s'∈* $\mathfrak{I}(S)$. *Then, for every* $\lambda \in L$ *the following properties hold:*

a) $C(s, 0) = s$,

b) $\lambda \leq \lambda' \Rightarrow C(s,\lambda) \supseteq C(s,\lambda')$,

c) $s \subseteq s' \Rightarrow C(s,\lambda) \subseteq C(s',\lambda)$

d) $C(s \cup s', \lambda) = C(s,\lambda) \cup C(s',\lambda)$ and $C(s \cap s', \lambda) = C(s,\lambda) \cap C(s',\lambda)$.

Given a fuzzy set $s: S \rightarrow L$ and a λ-cut $C(s, \lambda)$ of s, it is possible to define, for any $\lambda \in L$, another fuzzy set (see [38]),

$$\widetilde{C}_\lambda : C(s, \lambda) \rightarrow L$$

such that $\widetilde{C}_\lambda(x) = s(x)$. In this way, we obtain a particular class of fuzzy sets called *level fuzzy sets* of the fuzzy set s.

1.4 Fuzzy orderings

It is possible to generalize in a natural way the notion of crisp relation by the concept of *L-relation*, defined as an *L*-subset of a Cartesian product. More precisely, given two set S_1 and S_2, an *L- relation* from S_1 to S_2 is a map R: $S_1 \times S_2 \rightarrow L$. If L is the interval [0,1], the map R is called *fuzzy relation.* In this section we examine a particular class of binary *L*-relations, the *L- orderings* on a set *S*.

Let $*$ be a triangular norm and let ord: $S\times S \rightarrow L$ be an *L*- relation on *S*. Then we are interested to the following properties (see [22]) :

(1) $ord(x,x) = 1$ (reflexivity)

(2) $ord(x,y) = ord(y,x)$ (symmetry)

(3) $ord(x,y) * ord(y,z) \leq ord\,(x,z)$ ($*$- transitivity)

(4) $ord(x,y) = ord(y,x) = 1 \Rightarrow x = y$ (antisymmetry)

where $x, y, z \in S$.

Definition 1.4.1 An *L*- relation on *S* $ord : S\times S\rightarrow L$ is called:

- *L - preorder* if it satisfies **(1)** and **(3)** ,
- *L- order*, provided that it satisfies **(1)**, **(3)** and **(4)**,
- *L -similarity*, provided that it satisfies **(1)**, **(2)** and **(3)**
- *strict L- similarity*, provided that it satisfies **(1)**, **(2)**, **(3)** and **(4)**.

If *L* is the lattice [0,1] then we call these relations *fuzzy preorder, fuzzy order* and *fuzzy similarity,* respectively.

The *L*-preorders are also called *graded preference relations*, since *S* can be interpreted as a set of possible choices and $ord(x, y)$ as a measure of the preference of *y* with respect to *x*.

We say that *ord* is *crisp* provided that it assumes values only in the Boolean algebra {0, 1}. The notions proposed in Definition 1.4.1 extend the classical ones. It means that the crisp preorders (orders, similarities, strict similarities) coincide with the characteristic functions of the preorders (orders, equivalence relations, identity, respectively). By the λ-cuts we can establish a link between the L-orderings and the classical orderings.

Given an L-preorder *ord*, the cut $C(ord, 1)=\{(x, y)\in S\times S \mid ord(x, y)=1\}$ is always a preorder relation we denote by $\leq$ and we call the *preorder associated with ord*. In other words, $\leq$ is defined by setting $x \leq y$ if and only if $ord(x,y) = 1$. So, an L-preorder is an *L-order* if and only if $\leq$ is an order relation. Also, if *ord* is a similarity, then $C(ord,1)$ is an equivalence relation and if *ord* is a strict similarity, then $C(ord,1)$ is the identity relation. Let us remark that if $\lambda \neq 1$, then the closed λ-cut $C(ord, \lambda)=\{(x, y)\in S\times S \mid ord(x, y) \geq \lambda\}$ is not a preorder, in general.
Conversely, let $\leq$ be a preorder (order) relation, then its characteristic function is a L-preorder (L-order).

It is well known that any preorder $\leq$ on a set S induces an equivalence relation $\equiv$ defined by setting $x \equiv y$ provided that $x \leq y$ and $y \leq x$. In this way, in the quotient $S/\equiv$ we obtain an order relation by setting $[x]\leq[y]$ if $x \leq y$. Likewise, if *ord* is an L-preorder on S, considering the preorder $\leq$ associated with *ord*, we obtain an equivalence relation by setting

$$x \equiv y \iff x \leq y \quad \text{and} \quad y \leq x.$$

It means that x and y are equivalent if and only if $ord(x, y) = ord(y, x) = 1$ and we say that x is *similar* to y. Then once considered the quotient $S' = S/\equiv$, it is immediate to prove that the mapping

$$ord' : S'\times S' \to L \text{ such that } ord'([x], [y]) = ord(x, y)$$

is well defined and it is an L-order on S'. By this identification, it is always possible to change from an L-preorder relation to an L-order one. In particular, if *ord* is an L-similarity, on the quotient S' we still obtain an L-similarity ord' such that $ord'(x, y) = ord'(y, x) = 1 \Rightarrow x = y$.

Let us observe that, in literature, we can also find other methods useful to construct an L-order relation from an L-preorder relation. As an example (see [22]), let us consider a fuzzy preorder *ord* on a set S. It is possible to define a fuzzy similarity *eq* on S as $eq(x, y) = ord(x, y) \wedge ord(y, x)$. Then, by considering the equivalence relation $\equiv$ associated with *eq*, we obtain a fuzzy order *ord'* on the quotient $S/\equiv$, defined as $ord'([x], [y]) = ord(x, y)$.

We conclude this section by listing some useful propositions. First, let us recall that, given two sets equipped with two L-relations, (S, r) and (S', r'), a map $h : S \to S'$ is called *homomorphism* from (S, r) to (S', r') if

$$r(x, y) = r'(h(x), h(y)).$$

We say that h is an *isomorphism* if h is a one-to-one homomorphism.

Proposition 1.4.1 *If* (S, ord) *is an L-order, then any homomorphism defined in* (S, ord) *is injective.*

In the following proposition we show that any L-order on L induces an L-order on the class L^S of all the L-subsets of S.

Proposition 1.4.2 *Let* $ord : L\times L \to L$ *be an L-order on L whose associated order is the natural one on L, and define* $Incl : L^S \times L^S \to L$ *by setting*

$$Incl(s_1, s_2) = Inf\{ord(s_1(x), s_2(x)) \,/\, x\in S\}.$$

Then Incl is an L-order on L^S *whose associated order is the usual inclusion between L-subsets* (*see Definition 1.3.2*) .

Moreover,

Proposition 1.4.3 *Let L be a complete residuated lattice* (*see Definition 1.1.5*) *and S a nonempty set. Then the L-relation on L^S defined by setting:*

$$Incl(s_1, s_2) = Inf\{s_1(x) \rightarrow s_2(x) \ / \ x \in S\} \tag{1.5}$$

is an L-order whose associated order is the Zadeh inclusion between L-subsets.

Definition 1.4.2 Let L be a complete residuated lattice and S a nonempty set. Then we call *implication-based inclusion* the L-relation *Incl* defined by (1.5) and *implication-based inclusion space* any structure (C, *Incl*) where C is a class of L-subsets of S (see also [1]).

Let us observe that this definition is logic in nature. Indeed, in the first order multivalued logic, based on residuated lattices, the universal quantifier is interpreted by the operator *Inf*. And so, we can interpret the number $Incl(s_1, s_2)$ as the valuation of the claim "*for every x, if x belongs to s_1 then x belongs to s_2*".

1.5 Fuzzy similarities

The concept of similarity relation is essentially a generalization of an equivalence relation (see [51]). Moreover, by means of the notion of λ-cut, we can move from a similarity to a classical equivalence relation and vice versa. In the following, we make some observations referring to a similarity valued in [0,1].

As usual for a fuzzy set, also for a fuzzy similarity E on a set S, given $\lambda \in [0, 1]$, we can consider the closed (open) λ-cut $C(E, \lambda)=(x,y) \in S \times S \ / \ E(x,y) \geq \lambda\}$ ($O(E, \lambda)=\{(x,y) \in S \times S \ / \ E(x,y) > \lambda\}$, respectively). It can be proved that, if E is a fuzzy similarity, for every $\lambda \in [0,1]$, each λ- cut $C(E, \lambda)$ is an equivalence relation on S. More precisely, given $\lambda \in [0, 1]$, we obtain an equivalence relation R_λ by setting, $x \, R_\lambda \, y \Leftrightarrow (x, y) \in C(E, \lambda)$, for every $x, y \in S$. Conversely, if $\{R_\lambda \ / \ \lambda \in (0,1]\}$ is a nested family of distinct equivalence relations on S (i.e. $\alpha > \beta$ *iff* R_α

$\subset R_\beta$), then, for any choice of λ' s in (0,1] which includes $\lambda = 1$, $E = \cup_\lambda (\lambda, R_\lambda)$ is a similarity relation on S defined as

$$E(x, y) = sup\{min(\lambda,\ \chi_{R_\lambda}(x, y)) \ / \ \lambda \in (0,1]\}$$

where $min(\lambda,\ \chi_{R_\lambda}(x, y)) = \begin{cases} \lambda & if \ \ (x, y) \in R_\lambda \\ 0 & otherwise. \end{cases}$

In accordance, given a similarity E on S, we can consider the partition P_λ induced on S by $C(E, \lambda)$. If $\alpha \geq \beta$, then P_α is a refinement of P_β . Moreover, a similarity can be interpreted in terms of *fuzzy similarity classes* $E[x_j]$, defined, for every element x_j of the universe as $E[x_j]\ (x_i) = E\ (x_i\ , x_j)$, which is the grade of the membership of x_i in the fuzzy class $E[x_j]$.

Let us briefly recall that, given a first order language $\mathcal{L}$, we can define a *fuzzy model* for $\mathcal{L}$ as a pair (D, I) where, for any n-ary relation name r, $R = I(r)$ is a fuzzy subset of D^n, i.e. an n-ary fuzzy relation. Let us assume that the language $\mathcal{L}$ contains a binary relation name r and let us consider the following axioms, basic to define the notion of equivalence in classical set theory:

- $\forall x \ \ r(x, x)$,
- $\forall x \forall y \ (r(x, y) \rightarrow r(y, x))$,
- $\forall x \forall y \forall z \ (r(x, y) \wedge r(y, z) \rightarrow r(x, z))$;

It is evident that an interpretation $R = I(r)$ satisfies the above axioms if and only if the properties of reflexivity, symmetry and transitivity are satisfied, i.e. $R = I(r)$ is a fuzzy similarity. Let us observe that if we also consider the axiom $\forall x \forall y$ ($(r(x, y)) \wedge (r(y, x)) \rightarrow x = y)$ then the interpretation $R = I(r)$ satisfies the added axiom if and only if the property of antisymmetry (i.e. $R(x, y) = 1$ and $R(y, x) = 1 \Rightarrow x = y$)is satisfied.

1.6 Distances

In this section we give some information about structures consisting of a set equipped of a distance, which can be finite or infinite.

Let M be a non-empty set and d: $M \times M \to [0, \infty)$ be a mapping. Also, let us consider the following axioms for any $x, y, z \in M$:

(d1) $d(x, y) = 0 \Rightarrow x = y$,

(d'1) $d(x, x) = 0$, (*reflexivity*)

(d2) $d(x, y) = d(y, x)$, (*symmetry*)

(d3) $d(x, z) \leq d(x, y) + d(y, z)$, (*triangular inequality*)

(d'3) $d(x, z) \leq d(x, y) \vee d(y, z)$, (*strong triangular inequality*)

(d4) $d(x, y) = 0$ and $d(y, x) = 0 \Rightarrow x = y$.

Then (M, d) is called

- *metric space* if it satisfies (d1), (d'1), (d2) and (d3);
- *pseudometric space* if it satisfies (d'1), (d2) and (d3);
- *quasi-metric space* if it satisfies (d1), (d'1), (d3) and (d4);
- *quasi-pseudometric space* if it satisfies (d'1), (d3) and(d4);
- *semi- metric space* if it satisfies (d2), (d3) and (d4).

Likewise, if we have the axiom (d'3) instead of (d3) then (M, d) is called

- *ultrametric space,*
- *pseudo ultrametric space,*
- *quasi- ultrametric space,*
- *quasi - ultrapseudometric space,*
- *semi-ultrametric space,*

respectively. Finally, if the axiom (d4) is not required, then they are called *generalized spaces* (metric, ultrametric, pseudometric, etc.)

Then, as the word "*generalized*" refers to the lack of the axiom (d4), by relating to the usual definition of metric space:

- the word "*pseudo*" refers to the lack of the axiom (d1);
- the word "*quasi*" refers to the lack of the symmetric property (d2);
- the word "*semi*" refers to the lack of reflexivity;
- the word "*ultra*" refers to the fact that we consider the strong triangular inequality with the *maximum* operation $\vee$ instead of the sum.

Moreover, let us observe that (d'3) entails (d3). So, any ultrametric space is a metric space.

In the case that the map d takes values in the closed interval $[0, \infty]$ the spaces are called *extended*. In other words, the word *"extended"* indicates the possibility that a distance is infinite. In such a case, we can define the diameter of any subset X of M.

Definition 1.6.1 The *diameter* $D(X)$ of a subset X of M is the number in $[0,\infty]$ defined by setting

$$D(X) = Sup\{d(x,y) \,/\, x, y \in X\}.$$

If $D(X) \neq \infty$, then we say that X is *bounded*. We say that the space (M, d) is *bounded* provided that M is bounded.

Definition 1.6.2 Let (M',d') and (M, d) be two structures, where d': $M' \times M' \to [0,\infty]$ and d: $M \times M \to [0,\infty]$ are two distances. Then a map $i: M' \to M$ is called *isometry* provided that

$$d'(x, y) = d(i(x), i(y)),$$

for any $x, y \in M'$. An *isomorphism* is an one-to-one isometry.

1.7 Ultrametrics and quasi-metrics

Now we focus our attention to *ultrametric spaces* and to quasi-metric soaces

Ultrametric distances are not much known, but they are used in a lot of applications. Indeed, an ultrametric space can be represented by a *tree-structure* and so, these kind of distances are suitable for classification processes (see [29]).

Ultrametric spaces verify some interesting anomaluos properties, such as

(U1) If two open balls intersect, then a ball is included in the other one.

(U2) If two closed balls intersect, then a ball is included in the other one.

(U3) Every point in an open ball is a centre of the ball (egocentricity).

(U4) Every point in a closed ball is a centre of the ball (closed egocentricity).

(U5) Every open ball is closed and every closed ball is open.

(U6) Every triangle is isosceles and its base is less than or equal to the other two sides.

If we imagine an ultrametric space as having its points on a line or in a plane, we cannot appeal to our usual intuition for distance. Instead, it is useful to have a new framework for visualizing the ultrametric space, the tree-structure. In this way, it is easier to understand the properties. As an example, the fact that "triangles are always isoceles" is demonstrated by drawing the few different possible relative positions of three points, as in Figure 1. If two points q and r are close to one another, then their distances to a more distant point s must be the same.

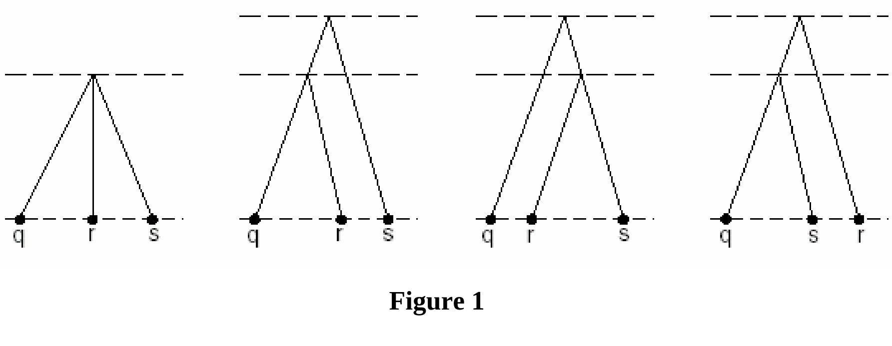


Figure 1

The picture of a ball is also simple. Given a point q and a distance γ, the set $\{x \,/\, d(x, q) \le \gamma\}$ is represented in an ultrametric tree by the set of all leaves in the subtree descending from a certain node (see Figure 2)

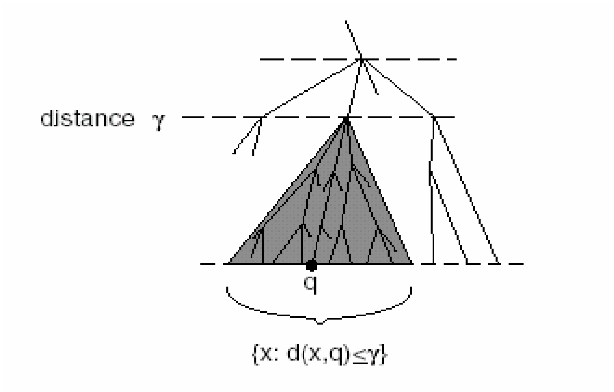


Figure 2

With this picture, it is easy to see why every point in a given ball is actually a centre of the ball. Let us suppose r is an arbitrary point in the ball of Figure 2. Then the ball centred at r, $\{x \,/\, d(x, r) \le \gamma\}$, is represented by the set of leaves in the subtree descending from the (unique) node above r at level γ. But this node is the same as that above q at level γ, giving the same ball.

Examples. Let $M \ne \varnothing$ and let d be the *discrete metric* on M defined by

$$d(x, y) = \begin{cases} 0 & \text{if } x = y \\ 1 & \text{if } x \ne y \end{cases}$$

Then (M, d) is an ultrametric space.

Let X be the set defined as $X = \{1/2^n \,/\, n \in \mathbb{N}\} \cup \{0\}$ and let d be a map defined by $d(x, y) = max\,\{x, y\}$ if $x \ne y$. Then (X, d) is an ultrametric space.

Now let us examine the second kind of distances covered in this section. We have seen in the previous section, that a quasi-metric on a set X is a distance function somewhat like a metric, but with significant weakening of the metric axioms. In particular, quasi-metrics are characterized by the lack of the symmetry axiom. We enunciate some interesting properties of this kind of distances and we furnish some examples of them.

Let us begin by observing that if d is a quasi-pseudometric on X, then the function d^{-1}, defined on $X \times X$ by $d^{-1}(x, y)=d(y, x)$, is still a quasi-pseudometric on X. Besides, we have

Proposition 1.7.1 *Any quasi-metric $d: X \times X \rightarrow \mathbb{R}^{+}$ is order-preserving with respect to the first variable and order-reversing with respect to the second variable.*

The quasi-metric spaces are related with the partial orders (see [10]) and we have:

Proposition 1.7.2 *Let (X, d) be a quasi-metric space, then the relation $\leq$ defined by setting, for any $x, y \in X$:*

$$x \leq y \Leftrightarrow d(x, y) = 0$$

is a partial order. Conversely, let $\leq$ be any partial order in a set X and let us define the map d: $X \times X \rightarrow \mathbb{R}^{+}$ by setting

$$d(x,y) = \begin{cases} 0 & \text{if } x \leq y \\ 1 & \textit{otherwise}. \end{cases}$$

Then (X, d) is a quasi-metric space whose associated partial order is $\leq$.

An interesting class of quasi-metric spaces is related to the *Hausdorff* distance. Indeed, given a metric space (M,δ), let $x \in M$ and X be a nonempty subset of M. We define $\delta(x, X)$ by setting $\delta(x, X) = Inf\{\delta(x, y) \mid y \in X\}$. Also, we can define the *excess function* e_δ by setting, for any X and Y in $P(M)$-$\{\emptyset\}$,

$$e_\delta(X, Y) = sup_{x\in X}\delta(x, Y).$$

The excess function results to be a quasi-metric.
Let us recall that the Hausdorff distance between two sets, X and Y, is defined by

$$\delta_H(X, Y) = max\{e_\delta (X, Y), e_\delta(Y, X)\}. \tag{1.6}$$

Let us give another example of quasi-metric distance (see [42]). Let $(D, \leq)$ be a Scott domain, i.e. an algebraic, bounded complete partial ordered set, and let B_D be the set of compact elements of D. Let us define a function r: $B_D \rightarrow \mathbb{N}$ such that $r^{-1}(n)$ is a finite set, for every n. Then the map D defined by

$$D(X, Y)=inf\{2^{-n} \mid Z\leq X \Rightarrow Z\leq Y \text{ for every } Z \text{ of rank } \leq n\} \tag{1.7}$$

is a quasi-metric.

Finally let us observe that, given a quasi-metric d on a set X, it is possible to define an associate metric d' on X, by setting $d'(x, y) = max\{d(x, y), d(y, x)\}$. We can call d' the symmetrization of d.

1.8 Some dualities between "closeness" and "distance"

There is an easy understandable duality between the notions of "closeness" and the one of "distance". Indeed, if one wants to compare some objects accordingly to their properties, it is possible to use or a measure of how they are "similar" or a measure of how they are "distant". Obviously, the smaller the distance is, the bigger the closeness is. More precisely, it is well-known that the notion of fuzzy similarity is connected to the non-fuzzy concept of pseudometric . The first notion seems more suitable than the second one in dealing with situations in which the objects involved do not verify sharply defined properties, but "vague" properties.

We start this section by analyzing the most simple case: the case in which the considered t-norm is the minimum t-norm.

The following proposition, whose proof is immediate, extends (to fuzzy orders and quasi-ultrametric distances) a connection between similarities and metrics proved in [45].

Proposition 1.8.1 *Let* $*$ *be the Gödel t-norm, let* $d: M \times M \to [0, 1]$ *be a map, and let us set*

$$ord\,(x, y) = 1 - d(x, y).$$

Then:

(i) *ord is a fuzzy similarity if and only if d is an pseudo ultrametric;*

(ii) *ord is a fuzzy preorder if and only if d is a generalized quasi ultrametric;*

(iii) *ord is a fuzzy order if and only if d is a quasi-ultrametric.*

As in the case of the fuzzy relations, if d is a generalized metric (ultrametric) space, the position

$$x \equiv y \quad \Leftrightarrow \quad d(x, y) = 0 \text{ and } d(y, x) = 0$$

defines an equivalence relation. Then it is possible to divide the space into the classes $[x] = \{y \in M \,/\, d(x, y) = d(y, x) = 0\}$. Moreover, it is immediate to prove that the mapping

$$d'\colon (M/\equiv) \times (M/\equiv) \to [0,1] \quad \text{such that} \quad d'([x], [y]) = d(x, y)$$

is well defined and it is a generalized metric (ultrametric) distance satisfying (d1) on the space of equivalence classes. By this identification it is possible to change from a pseudometric structure to a metric one. So, from Proposition 1.8.1 it follows that *ord'* is a similarity if and only if d' is an ultrametric distance.

The fuzzy similarity with the Gödel t-norm is a limited notion because it is equivalent to a restricted class of metrics, the class of ultrametrics. For similarities with a t-norm different from the Gödel one it is possible to obtain an analogous result to Proposition 1.8.1. To this aim, it is useful the notion of additive generator of a t-norm and the notion of its pseudoinverse (Definition 1.2.5). By means of the Archimedean t-norms' characterization we are able to extend a connection between pseudometrics and fuzzy-similarities exposed in [45].

Proposition 1.8.2 *Let* $*$ *be a continuous Archimedean t-norm and* f: $[0, 1] \to [0, \infty]$ *an additive generator of* $*$. *Moreover, let* d: $M\times M \to [0, 1]$ *be a map and define the fuzzy relation* $ord_f(d)$: $M\times M \to [0, 1]$ *by setting*

$$ord_f(d)(x, y) = f^{[-1]}(d(x,y)).$$

Then:

(i) *d is an extended (generalized) pseudometric* $\Rightarrow$ *$ord_f(d)$ is a $*$-fuzzy similarity*

(ii) *d is an extended quasi-metric* $\Rightarrow$ *$ord_f(d)$ is a $*$-fuzzy order.*

(iii) *d is an extended generalized quasi-metric* $\Rightarrow$ *$ord_f(d)$ is a $*$-fuzzy preorder*;

(iv) *d is an extended metric* $\Rightarrow$ *$ord_f(d)$ is a strict $*$-fuzzy similarity.*

Proof. (i) The reflexivity is immediate, since $ord_f(d)(x,x) = f^{[-1]}(d(x,x))= f^{[-1]}(0) = 1$. The symmetry follows from definitions, trivially. To prove the $*$ -transitivity, i.e.

$$ord_f(d)\,(x, y) * ord_f(d)\,(y, z) \leq ord_f(d)\,(x, z) \qquad (1.6)$$

let us observe that, by the definition of pseudoinverse, in the case $d(x,y)\notin f([0,1])$ we have that $ord_f(d)(x, y)= f^{[-1]}(d(x,y)) = 0$ and in the case $d(y,z)\notin f([0,1])$, we have that $ord_f(d)(y, z) = f^{[-1]}(d(y,z)) = 0$. In both the cases, (1.6) is trivial. If we take $x, y, z \in M$ such that both $d(x,y)$ and $d(y,z) \in f([0,1])$, then, by (1.4), we have

$$\begin{aligned} ord_f(d)\,(x, y) * ord_f(d)\,(y, z) &= f^{[-1]}(d(x,y)) * f^{[-1]}(d(y,z))= \\ &= f^{[-1]}(\, f(f^{[-1]}(d(x,y))\,) + f(f^{[-1]}(d(y,z))\,)\,) = \\ &= f^{[-1]}(\, d(x,y) + d(y,z)\,) \leq f^{[-1]}(d(x, z)) = ord_f(d)(x, z), \end{aligned}$$

because $f^{[-1]}$ is strictly decreasing.

(ii) We have to prove the antisymmetry of $ord_f(d)$. So, let x, $y \in M$ such that $ord_f(d)(x, y) = 1 = ord_f(d)(y, x)$. From this condition follows that $f^{[-1]}(d(x, y)) = 1 = f^{[-1]}(d(y, x))$, and therefore $f^{-1}(d(x,y)) = 1 = f^{-1}(d(y,x))$. Then $d(x,y) = 0 = d(y,x)$ and, by the antisymmetry of d, $x = y$.

(iii) (iv) The proofs are analogue to the previous ones.

□

Examples. Let d be the usual distance in an Euclidean space and let $f(x) = 1-x$. Then $ord_f(d)(x,y) = 1-d(x,y)$ if $d(x,y) \leq 1$, and $ord_f(d)(x,y) = 0$ otherwise. $ord_f(d)$ is a $*$-fuzzy order where $*$ is the Lukasiewicz norm.

As another example, let us assume that $f(x) = -log(x)$. Therefore, we set $ord_f(d)(x,y) = e^{-d(x,y)}$ and we obtain a $*$-fuzzy order, where $*$ is the product t-norm. As a matter of fact, in both the examples $ord_f(d)$ is a strict $*$-similarity.

Conversely, we can associate any fuzzy order with an extended metric.

Proposition 1.8.3 *Let f:[0, 1] $\rightarrow$ [0, ∞] be an additive generator and $*$ be the related t-norm. Let ord: $M \times M \rightarrow [0, 1]$ be a map and consider the function $d_f(ord)$: $M \times M \rightarrow [0, \infty]$ defined by setting*

$$d_f(ord)(x, y) = f(ord(x, y))$$

Then:

(i') *ord is a $*$- fuzzy similarity $\Rightarrow$ $d_f(ord)$ is an extended generalized pseudometric;*

(ii') *ord is a $*$- fuzzy order $\Rightarrow$ $d_f(ord)$ is an extended quasi-pseudometric;*

(iii')*ord is a $*$- fuzzy preorder $\Rightarrow$ $d_f(ord)$ is an extended generalized quasi-pseudometric;*

(iv') *ord is a strict $*$- fuzzy similarity $\Rightarrow$ $d_f(ord)$ is an extended metric.*

Proof. (i') For any $x \in M$, $d_f(ord)(x,x) = f(ord(x,x)) = f(1) = 0$ and so the reflexivity is proved The symmetry of $d_f(ord)$ follows from definitions, immediately. Before proving the triangular inequality of $d_f(ord)$, i.e.

$$f(ord(x,y)) + f(ord(y,z)) \geq f(ord(x,z)),$$

we recall that

$$f(f^{[-1]}(x)) = \begin{cases} x & \text{if } x \in f([0, 1] \\ f(0) & \text{otherwise} \end{cases}$$

where $f(0)$ is the maximum of the function. Then, since f is decreasing and ord is $*$- transitive, we observe that,

$f(\ ord(x, y)*ord(y, z)\ ) \geq f(ord(x, z))$ and therefore, by (1.4),

$f(\ f^{[-1]}(\ f(ord(x, y)) + f(ord(y, z))\ )) \geq f(ord(x, z))$.

Now, if $f(ord(x, y)) + f(ord(y, z)) \in f([0,1])$, we obtain that

$f(ord(x, y)) + f(ord(y, z)) \geq f(ord(x, z))$.

Otherwise,

$f(ord(x, y)) + f(ord(y, z)) \geq f(0) \geq f(ord(x, z))$.

(ii') We have to prove the antisymmetry of $d_f(ord)$. Let $x, y \in M$ such that $d_f(ord)(x, y) = 0$ and $d_f(ord)(y, x) = 0$. Then $f(ord(x,y)) = 0 = f(ord\ (y,x))$, and hence $ord(x,y) = 1 = ord(y,x)$. From the antisymmetry of ord it follows that $x = y$.

(iii') (iv) The proofs are analogue to the previous ones.

□

Let us provide in the following table some examples.

T-NORM	**ADDITIVE GENERATOR**	**DISTANCE**
Product $a*b = a \cdot b$	$f(x) = -log(x)$	$d_f(ord)(x,y) = -ln\ (ord(x,y))$

Lukasiewicz $a*b=max(0, a+b-1)$	$f(x)= 1-x$	$d_f(ord)(x,y)= 1- ord(x,y)$

Table 1

The established connection in the last two propositions is not completely satisfactory, in a sense. In the next proposition we observe that, while $ord_f(d_f(ord))= ord$, it results $d_f(ord_f(d)) \neq d$, in general. Indeed, we have the following.

Proposition 1.8.4 *Let f be an additive generator of a t-norm* $*$. *Then, for any fuzzy preorder ord,*

$$ord_f(d_f(ord))= ord.$$

Moreover, for any extended generalized quasi-metric $d : S\times S \rightarrow [0, \infty]$, *we have*

$$d_f(ord_f(d))= d\wedge f(0).$$

Proof. Observe that $f(f^{[-1]}(d(x,y))) = d(x,y)$ if $d(x,y)\leq f(0)$ and $f(f^{[-1]}(d(x,y))) = f(0)$ otherwise.

□

Then, given an additive generator f, the resulting connection among $*$-fuzzy preorders and extended generalized quasi-metrics works well only for the extended generalized quasi-metrics (M, d) such that the diameter $D(M)\leq f(0)$. As an example, if f coincides with $-log$, then since $f(0) = \infty$, all works well. Instead, if $f(x) = 1-x$, and d is the usual Euclidean distance, then $d_f(ord_f(d))(x,y) = d(x,y)$ if $d(x,y) \leq 1$ and $d_f(ord_f(d))\ (x,y) = 1$ otherwise.

Chapter 2

Incomplete and fuzzy information spaces

The dualities examined in Chapter Chapter 1 allow us to establish a link between some "metric" structures, in the context of point-free geometry, and some structures equipped with fuzzy relations. These notions are defined on suitable spaces of "regions".

2.1 Point-free geometry and incomplete pieces of information

The aim of point-free geometry is to give an axiomatic basis to geometry in which the notion of *point* is not assumed as a primitive. In this direction, geometry can be built up by assuming as primitive the notions of region or solid and, thereafter defining the points in a suitable way. If we want to refer to pointless geometry in terms of the vocabulary of logic, we say that regions are considered as individuals, i.e., first order objects, while points are represented by classes (or sequences), i.e. second order objects.

Several authors addressed their researches to attempts of building a geometry "without points". One of the first example in such a direction was furnished by Whitehead's researches, where the primitive notions are the *regions* and the *inclusion* relation between regions (see [46], [48]). Anyhow, this approach seems more suitable as a basis for a *"mereology"*, i.e. an investigation about the set theoretical *part-whole* relation, rather than about a point-free geometry. So, it is not surprising the fact that, later, Whitehead proposed a different approach, topological in nature, in which the primitives are the *regions* and the *connection relation*, that is, the relation between two regions that either overlap or have at least a common boundary point (see [47]).

Recently the increased interest in point-free geometry is due to the different reasons. As an example, one of them is related to the complexity, from a computational point of view, of the Euclidean geometry based on the notion of point. Our interest in the point-free approach mainly derived from the adequacy of the notion of *region* to represent incomplete information. Indeed, given a region, the measure of its *diameter* can be interpreted as a measure of the incompleteness of the available information: the bigger the diameter is, the less complete the information is. *Points*, having zero-diameter, represent precise information.

2.2 Metrical approach to point-free geometry

In accordance with Whitehead ideas, Gerla proposed in [23], [24] a system of axioms for the pointless spaces theory in which *regions, inclusion, distance* and *diameter* are assumed as primitives and, in order to give a metric approach to point-free geometry, defined the notion of *pointless metric space*.

Definition 2.2.1 A *pointless pseudometric space*, briefly *ppm-space*, is a structure $(R, \leq, \delta, |\,|)$, where

• $(R, \leq)$ is an ordered set;

• $\delta : R\times R\rightarrow[0,\infty)$ is an order-reversing map, i.e. $x \geq y \Rightarrow \delta(y, z) \leq \delta(x, z)$, for every $x, y, z \in R$;

• $|\,|:R \rightarrow[0, \infty]$ is an order-preserving map, i.e. $x \geq y \Rightarrow |x| \geq |y|$ for every $x, y \in R$

and, for every $x, y, z \in R$ the following axioms hold:

(a1) $\delta(x, x) = 0$

(a2) $\delta(x, y) = \delta(y, x)$

(a3) $\delta(x, y) \leq \delta(x, z) + \delta(z, y) + |z|$ (*generalized triangle inequality*).

The elements in R are called *regions*, the relation $\leq$ *inclusion*, the number $\delta(x, y)$ the *distance* between the regions x and y and the number $|x|$ the *diameter* of x. A region x is *bounded* if its diameter $|x|$ is finite.

Definition 2.2.2 We call *atoms* the minimal elements of R, with diameter equal to zero.

Let us observe that *ppm*-spaces generalize pseudometric spaces; indeed pseudometric spaces coincide with the *ppm*- spaces for which every region is an atom, that is the order relation coincides with the identity relation and $|\ |$ is constantly equal to zero.

To define *pointless metric spaces*, let us recall that a metric space (M, d) is a pseudometric space such that $x=y \Leftrightarrow d(x, y)= 0$; in other words the identity relation can be defined via the distance function. In accordance, we give the following

Definition 2.2.3 A *pointless metric space*, briefly *pm-space*, is a *ppm*-space $(R, \leq, \delta, |\ |)$, such that

$$x \geq y \Leftrightarrow |x| \geq |y| \text{ and } \delta(x, z) \leq \delta(y, z) \text{ for every } z \in R..$$

As we will show in Proposition 2.5.1, this equivalence supports a way to define the inclusion from a distance and a diameter.
Metric spaces coincide with the *pm*-spaces such that all the regions are atoms.

We introduce now a particular class of pointless metric spaces, related with the notion of ultrametric spaces: the class of *pointless ultrametric spaces* (see [20]).

Definition 2.2.4 A *pointless ultrametric space*, briefly *pu*-space, is a *pm*-space $(R, \leq, \delta, |\ |)$ such that

(A3) $\delta(x, y) \leq \delta(x, z) \vee \delta(z, y) \vee |z|$ (*generalized strong triangle inequality*),
where $\vee$ is the maximum.

Let us observe that, since

$$\delta(x, z) \vee \delta(z, y) \vee |z| \le \delta(x, z) + \delta(z, y) + |z|,$$

then (A3) entails (a3).

In a *ppm*-space it is possible to define *points* by means of a procedure similar to the completion of a metric space by Cauchy sequences. Let us describe such a procedure by the introduction of the notion of *abstraction process*.

Definition 2.2.5 An *abstraction process* of a *ppm*-space R is a sequence $\langle p_n \rangle_{n\in\mathbb{N}}$ of nonempty bounded regions such that

a) $lim_{n\to\infty} | p_n | = 0$,

b) $\forall \varepsilon > 0\ \exists \nu \in \mathbb{N}$ *such that* $\forall h, k \ge \nu$, $\delta(p_h, p_k) < \varepsilon$.

We denote by $AP(R)$ the class of the abstraction processes of R.

Decreasing sequences of nonempty regions with vanishing diameters are examples of abstraction processes. It is possible that in a *ppm*-space there is no abstraction process. Let us enunciate the following proposition, whose proof we omit.

Proposition 2.2.1 *Let* $(R, \le, \delta, |\ |)$ *be a ppm-space and let* $AP(R)$ *be nonempty.Let us define the map dis:* $AP(R) \times AP(R) \to [0, \infty)$ *by*

$$dis(\langle p_n \rangle_{n\in\mathbb{N}}, \langle q_n \rangle_{n\in\mathbb{N}}) = lim_{n\to\infty}\delta(p_n, q_n),$$

for every $\langle p_n \rangle_{n\in\mathbb{N}}, \langle q_n \rangle_{n\in\mathbb{N}} \in AP(R)$. *Then* $(AP(R), dis)$ *is a pseudometric space.*

We denote by $(M(R), dis)$ the metric space obtained as a quotient of $(AP(R), dis)$, modulo the relation $\equiv$ defined by $\langle p_n \rangle \equiv \langle q_n \rangle \Leftrightarrow dis(\langle p_n \rangle, \langle q_n \rangle) = 0$. In accordance, we call *point* every element of $M(R)$, i.e. a point P is a class

$$[\langle p_n \rangle] = \{\langle q_n \rangle \in AP(R) \ /\ \langle q_n \rangle \equiv \langle p_n \rangle\}.$$

The distance between points *dis*: $M(R)\times M(R)\rightarrow [0, \infty)$ is defined by setting, for every $P, Q \in M(R)$

$$dis(P, Q) = dis(\langle p_n \rangle_{n\in\mathbb{N}}, \langle q_n \rangle_{n\in\mathbb{N}}) = lim_{n\to\infty}\delta(p_n, q_n),$$

where $\langle p_n \rangle, \langle q_n \rangle \in AP(R)$ are representatives of P and Q, respectively.

2.3 Canonical *pm*-spaces

A class of basic examples of *pm*-spaces and *pu-spaces* is obtained by starting from a class of subsets of a pseudometric space, with the usual inclusion relations $\subseteq$ between subsets.

Proposition 2.3.1 *Let (M, d) be a pseudometric space and let C be a nonempty class of bounded and nonempty subsets of M. Define δ and $|\ |$ by setting*

$$\delta(x, y) = inf\{d(X,Y) \ /\ X\in x,\ Y\in y\}$$

and

$$|x|= sup\{d(X, Y) \ /\ X, Y\in x\},$$

for every $x, y \in C$. Then $(C, \subseteq, \delta, |\ |)$ is a pm-space. Besides, if (M, d) is a pseudoultrametric space, then $(C, \subseteq, \delta, |\ |)$ is a pu-space.

Proof. (a1) and (a2) are immediate. To prove (a3), let x, y and z be subsets of M, $X\in x$, $Y\in y$, Z and $Z'\in z$; then

$$\delta(x, y) \le d(X,Y) \le d(X, Z) + d(Z, Z') + d(Z', Y) \le d(X, Z) + d(Z', Y) + |z|.$$

Consequently,

$$\delta(x, y) \le \delta(x, z) + \delta(z, y) + |z|.$$

Now assume that (M, d) is a pseudoultrametric space. Then

$$\delta(x, y) \le d(X, Y) \le d(X, Z) \vee d(Z, Z') \vee d(Z',Y) \le d(X, Z) \vee d(Z', Y) \vee |z|,$$

and therefore $(C, \subseteq, \delta, |\ |)$ is a *pu*-space.

□

The so obtained spaces are called *canonical*.

Thanks to these models, it is possible to clarify the meaning of the generalized triangular inequality (axiom (a3) of Definition 2.2.1). Let us observe Figure 3:

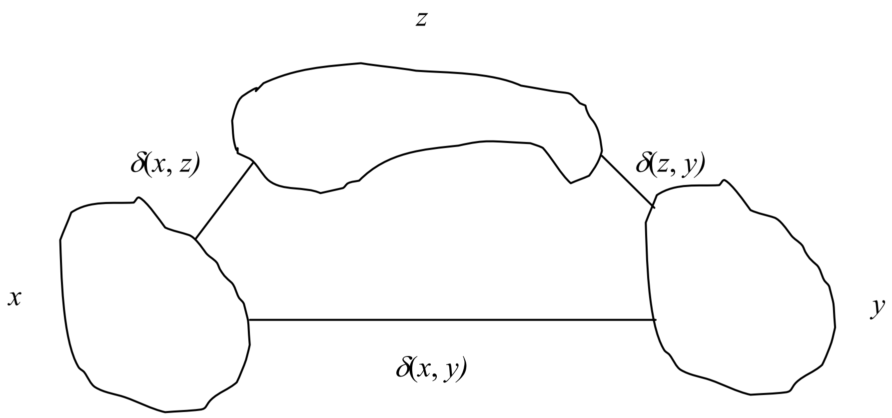


Figure 3

We can guess that in such a case $\delta(x, y) \geq \delta(x, z) + \delta(z, y)$, so it is necessary to consider $|z|$. In other words, it is necessary to take in account the incompleteness of the information represented by the region z, more precisely, by the measure of its diameter.

2.4 Semimetrics and semisimilarities

In this section we introduce two new classes of structures. The structures in the first one are based on distances which satisfy symmetry and a triangular inequality, whereas reflexivity is not required. In other words, we consider semimetrics in the context of point-free geometry. The second class is made up of structures with a particular fuzzy relation. We find also a connection between these two classes, which result to be dual.

Definition 2.4.1 A *semi-metric space,* briefly *sm-space*, is a structure (R, d) where R is a set whose elements are called *regions* and $d: R \times R \to [0, \infty]$ is a function we call *semi-distance*, verifying, for any $x, y, z \in R$ axioms (d2) and (d3) of Section 1.6:

(d2) $d(x, y) = d(y, x)$,

(d3) $d(x, y) \leq d(x, z) + d(z, y)$.

Let us remark that the condition $d(x, y)=0$ is not required. Given a semi-distance d, we define a *diameter* by setting:

$$|x|_d = d(x, x) \ . \tag{2.1}$$

Let us observe that by setting $y = x$ and $z = y$ in (d3), we obtain that

$$d(x, x) \leq d(x, y) + d(y, x)$$

and therefore, by (d2), that $d(x, x) \leq 2d(x, y)$. Likewise we have that $d(y, y) \leq 2d(x, y)$ and therefore it results

$$d(x, y) \geq \frac{|x|_d}{2} \vee \frac{|y|_d}{2} .$$

So we can have $d(x, y) = 0$ only in the case both x and y have zero-diameter.

As we saw in Section 1.8 of Chapter Chapter 1, in literature it is possible to find a duality between the notion of metric and the notion of similarity (for example in [26]). Likewise we can give the next definition as a dual concept of semidistance.

Definition 2.4.2 Let $*$ be a t-norm. A *semisimilarity* is a fuzzy relation E on R such that

(e1) $E(x, y) = E(y, x)$ (symmetry)

(e2) $E(x, z) * E(z, y) \leq E(x, y)$ (transitivity)

for every $x, y, z \in R$.

Let us recall that a *similarity* is a semisimilarity such that

(e3) $E(x, x) = 1$.

$E(x, y)$ is regarded as truth-value of a statement like $x =_R y$. Semisimilarities are used to give a general approach to fuzzy sets theory based on the notion of category (see also M. Fourman and D.S. Scott [18]), and it is possible to show that they are strictly related with *sm*-spaces. Two cases regarding Definition 2.4.2 are examined: the case of Archimedean t-norms and the case of the *Gödel* t-norm. In the first one we use, (as in Gerla, [25]), the connection examined in Section 1.8 of Chapter Chapter 1. We have:

Proposition 2.4.1 *Let f:* $[0, 1]\to[0, \infty]$ *be an additive generator of an Archimedean t-norm* $*$ *and let d be a semidistance on a set R..Then the fuzzy-relation* $E_f(d)$ *defined by*

$$E_f(d)(x, y) = f^{[-1]}(d(x, y))$$

is a semisimilarity with respect to the t-norm $*$.
Conversely, let E be a semisimilarity on R with respect to the t-norm $*$, *then the structure* $(R, d_f(E))$ *where* $d_f(E)$ *is defined by*

$$d_f(E)(x, y) = f(E(x, y)),$$

is a sm-space.

Proof. It is analogous to that one of Proposition 1.8.2 and Proposition 1.8.3.

If the t-norm is the *Gödel* t-norm the transitivity becomes
(e2*) $E(x, z) \wedge E(z, y) \leq E(x, y)$.
In such a case, we call the semisimilarity *G-semisimilarity* and, setting $y = x$ in (e2*), we obtain that

$$E(x, z) \wedge E(z, x) \leq E(x, x)$$

and therefore that $E(x, z) \leq E(x, x)$. Then

$$E(x, z) \leq E(x, x) \wedge E(z, z).$$

Since the *Gödel* t-norm is not Archimedean, Proposition 2.4.1 doesn't hold for it. So, in this case, we consider a subclass of *sm*-spaces, shrinking the codomain of the semidistance and adding an axiom.

Definition 2.4.3 A *semi-ultrametric space*, briefly *su-space*, is a *sm*-space (R, d), where the semi-distance is a function $d: R\times R\rightarrow[0, 1]$, such that, for any $x, y, z \in R$:

(d'3) $d(x, y) \leq d(x, z) \vee d(z, y)$.

Obviously, (d'3) entails (d3). Let us observe that by setting $y = x$ and $z = y$ in (d'3), we obtain that $d(x, x)\leq d(x, y) \vee d(y, z)$ and therefore, by (d2), that $d(x, x)\leq d(x, y)$. Likewise we have that $d(y, y)\leq d(x, y)$ and therefore it results

$$d(x, y) \geq |x|_d \vee |y|_d.$$

Now we are able to describe the relation between the *G*-semi-similarities and the *su*-spaces.

Proposition 2.4.2 *Let d be a semi-ultrametric on a set R, then the fuzzy-relation E_d defined by*

$$E_d(x, y) = 1-d(x, y) \tag{2.2}$$

is a G-semi-similarity. Conversely, let E be a G-semisimilarity on R, then the structure (R, d_E), defined by

$$d_E(x, y) = 1-E(x, y) \tag{2.3}$$

is a su-space.

Proof. Let E_d be defined by (2.2). Then (e1) is immediate. To prove (e2*) observe that

$$E_d(x, y) \wedge E_d(y, z) = (1- d(x, y)) \wedge (1- d(y, z))$$
$$= 1-(d(x, y) \wedge d(y, z))\leq 1- d(x, z) = E_d(x, z).$$

Now let us consider d_E defined by (2.3). Then (d2) is immediate. To prove (d'3) it is sufficient to observe that

$$d(x, y) = 1 - E(x, y) \leq (1-E(x, z)) \vee (1-E(z, y))$$
$$= d(x, z) \vee d(z, y).$$

□

Now we give a characterization of G-semisimilarities in terms of related cuts and by means of the notion of semiequivalence.

Definition 2.4.4 Let S be a nonempty set. A (classical) relation R on S is called *semiequivalence* provided that it is symmetric and transitive.

Let us denote by $D_R = \{x \in S$ / there is an element $y \in S : (x, y) \in R\}$ the domain of R. Then,

Proposition 2.4.3 *A relation R on a set S is a semiequivalence if and only if is symmetric and it is an equivalence relation on its domain.*

Proof. Let R be a semi-equivalence relation. It results that if $x \in D_R$, then $(x, x) \in R$, that is, R is reflexive in D_R. Indeed, if y is such that $(x, y) \in R$, by the symmetry $(y, x) \in R$ and, in account of the tranisitivity, we have $(x, x) \in R$. Therefore, every semiequivalence relation R on S is an equivalence relation on its domain D_R.

Vice versa, if R is an equivalence relation on D_R and if it is symmetric on S, then R is a semiequivalence relation on S. Indeed let $x, y, z \in S$ such that $(x, y) \in R$ and $(y, z) \in R$. By the symmetry $(z, y) \in R$ and by the definition of D_R it results that $x, y, z \in D_R$. By the transitivity on D_R it results $(x, z) \in R$.

□

Definition 2.4.5 A family $(R_\lambda)_{\lambda\in[0,1]}$ of semiequivalence relations on a set S is called *order-reversing* if it results that

- $R_\beta \subseteq R_\alpha$ for every $\alpha \le \beta$, $\alpha, \beta \in [0,1]$;
- $R_0 = S \times S$.

Proposition 2.4.4. *A fuzzy relation E is a G-semisimilarity if and only if the cuts of E define an order-reversing family* $(C(E, \lambda))_{\lambda\in[0,1]}$ *of semiequivalences.*

Also, any order-reversing family of semiequivalence relations defines a *G*-semisimilarity.

Proposition 2.4.5 *Let* $(R_\lambda)_{\lambda\in[0,1]}$ *be an order-reversing family of semiequivalence relations. Then the fuzzy relation E defined by setting*

$$E(x, y) = Sup\{\lambda /(x, y) \in R_\lambda\}$$

is a G-semisimilarity.

Proof. Condition (e1) is immediate by the symmetry of R_λ. To prove (e2*), let us consider

$E(x, z) = Sup\{\lambda / (x, z) \in R_\lambda \} = \mu$

$E(z, y) = Sup\{\lambda / (z, y) \in R_\lambda \} = \xi$

$E(x, y) = Sup\{\lambda / (x, y) \in R_\lambda \} = \eta$.

Let us suppose $\mu \le \xi$ (likewise $\xi \le \mu$). Since $(R_\lambda)_{\lambda\in[0,1]}$ is an order-reversing family of relations, it results $R_\xi \subseteq R_\mu$. Therefore we have $(x, z) \in R_\mu$ and $(z, y) \in R_\mu$ and then, by transitivity, $(x, y) \in R_\mu$. But $\eta = Sup\{\lambda / (x, y) \in R_\lambda \}$, then $\eta \ge \mu$ and, since $\mu \wedge \xi = \mu$, the condition (e2*)

$$E(x, z) \wedge E(z, y) \le E(x, y)$$

is verified.

□

2.5 Connections between pointless metric spaces and semimetric spaces

In order to establish a connection between *pm*-spaces and *sm*-spaces, we observe that in defining *pm*-spaces we can consider the inclusion relation as a derived notion. In fact, as proved in [24], the following holds true:

Proposition 2.5.1 *Let* $(R,\ \delta,\ |\ |)$ *be a structure satisfying* (a1), (a2) *and* (a3) (*of Definition 2.2.1*) *and let us define* $\leq$ *by setting*

$$x \leq y \ \textit{ iff } \ |x| \leq |y| \textit{ and } \delta(x, z) \geq \delta(y, z),$$

for any $z \in R$.. *Then* $(R,\ \delta,\ |\ |)$ *is a pm-space.*

Proof. We observe only that Definition 2.2.3 is trivially satisfied.

□

In accordance with such a proposition, in the following we denote by (R, δ, $|\ |$) a *pm*-space whose order relation is defined as in Proposition 2.5.1.

Now let us see how it is possible to associate any *pm*-space with a *sm*-space.

Proposition 2.5.2 *Let* $(R,\ \delta,\ |\ |)$ *be a pm-space and let* $d_\delta{:}R\times R\to[0,\ 1]$ *be defined by setting, for any* $x, y \in R$,

$$d_\delta\,(x, y) = \delta(x, y) + \frac{|x|}{2} + \frac{|y|}{2}. \tag{2.4}$$

Then the structure $(R,\ d_\delta)$ *is a sm-space whose diameter coincides with* $|\ |$.

Proof. (d2) and the equality $|\ |_d = |\ |$ are trivial. Besides,

$$d_\delta\,(x, y) = \delta(x, y) + \frac{|x|}{2} + \frac{|y|}{2} \leq \delta(x, z) + \delta(z, y) + |z| + \frac{|x|}{2} + \frac{|y|}{2}$$

$$= \left(\delta(x, z) + \frac{|x|}{2} + \frac{|z|}{2}\right) + \left(\delta(z, y) + \frac{|z|}{2} + \frac{|y|}{2}\right)$$

$$= d_\delta(x, z) + d_\delta(z, y).$$

□

Conversely, it is possible to associate any *sm*-space with a *pm*-space.

Proposition 2.5.3 *Let* (R, d) *be a sm-space and let* $\delta_d{:}R\times R\to[0, 1]$ *be defined by setting*

$$\delta_d(x, y)= \begin{cases} d(x,y)-\dfrac{|x|_d}{2}-\dfrac{|y|_d}{2} & if\ d(x,y)\geq\dfrac{|x|_d}{2}+\dfrac{|y|_d}{2} \\ 0 & otherwise \end{cases}$$

Then the structure $(R, \delta_d, |\ |_d)$*, where* $|\ |_d$ *is defined by* (2.1)*, is a pm-space .*

Proof. Axioms (a1) and (a2) are immediate. If $d(x, y)< \dfrac{|x|_d}{2}+\dfrac{|y|_d}{2}$, then (a3) is trivial. Otherwise, by (d3),

$$\delta_d(x, y) = d(x, y) - \frac{|x|_d}{2} - \frac{|y|_d}{2} \leq d(x, z) + d(z, y) - \frac{|x|_d}{2} - \frac{|y|_d}{2} + \left(|z|_d - \frac{|z|_d}{2} - \frac{|z|_d}{2}\right)$$

$$= \delta_d(x, z) + \delta_d(z, y) + |z|_d\,.$$

□

Let us observe that the definitions of d_δ and δ_d in Proposition 2.5.2 and in Proposition 2.5.3 are not the unique possible ways to associate a *pm*-space with a *sm*-space and vice versa. For example, it is possible to associate any *pu*-space $(R, \delta, |\ |)$ with a *su*-space (R, d_δ) by setting

$$d_\delta (x, y) = \delta(x, y) \vee |x| \vee |y|, \tag{2.5}$$

for any $x, y \in R$. Thus, we have:

Proposition 2.5.4 *Let* $(R, \delta, |\ |)$ *be a pu-space, then the structure* (R, d_δ) *defined by* (*2.5*) *is a su-space whose diameter coincides with* $|\ |$.

Proof. (d2) and the equality $|\ |_d = |\ |$ are trivial. Besides,

$$d_\delta(x, z) = \delta(x, z) \vee |x| \vee |z| \leq \delta(x, y) \vee \delta(y, z) \vee |y| \vee |x| \vee |z|$$

$$= (\delta(x, y) \vee |x| \vee |y|) \vee (\delta(y, z) \vee |y| \vee |z|) = d_\delta(x, y) \vee d_\delta(y, z).$$

□

Conversely, we can associate any *su*-space (R, d) with a *pu*-space $(R, \delta_d, |\ |_d)$ by setting, for any $x, y \in R$,

$$\delta_d(x, y) = \begin{cases} d(x, y) & \text{if } d(x, y) = |x|_d \vee |y|_d \\ 0 & \text{if } d(x, y) > |x|_d \vee |y|_d \end{cases} \tag{2.6}$$

In such a way, we obtain:

Proposition 2.5.5 *Let* (R, d) *be a su-space, then the structure* $(R, \delta_d, |\ |_d)$ *defined by* (2.6) *and* (2.1) *is a pu-space .*

Proof. Axioms (a1) and (a2) are immediate. To prove that

$$\delta_d(x, z) \vee \delta_d(z, y) \vee |z|_d \geq \delta_d(x, y)$$

let us assume that $d(x, z) \geq d(z, y)$. Now, in the case $\delta_d(x, y) = 0$ and in the case $|z|_d \geq d(x, y)$ such an inequality is trivial. So, it is not restrictive to assume that $d(x, y) > |x|_d \vee |y|_d$, and therefore that $\delta_d(x, y) = d(x, y)$ and that $d(x, y) > |z|_d$. In such a case, by (d'3), we have

$$d(x, z) = d(x, z) \vee d(z, y) \geq d(x, y) > |x|_d \vee |z|_d$$

and therefore

$$\delta_d(x, z) \vee \delta_d(z, y) \geq \delta_d(x, z) = d(x, z) \vee d(z, y) \geq d(x, y) = \delta_d(x, y).$$

Likewise we proceed in the case $d(x, z) \leq d(z, y)$.

□

Besides, we can also set

$$\delta_d(x, y)= \begin{cases} d(x,y) & \text{if } d(x,y) > |x|_d \vee |y|_d \\ 0 & \text{if } x = y \end{cases}$$

instead of (2.6).

2.6 Pointless ultrametric spaces and the category of fuzzy sets

In order to establish a link between point-free geometry and the categorical approach to fuzzy sets theory proposed by Höhle in [28] we propose a direct connection between *pu*-spaces and structures with *G*-semisimilarity.

In accordance with Proposition 2.4.2, Proposition 2.5.4 and Proposition 2.5.5, any connection between *su*-spaces and *pu*-spaces is also a connection between structures with semisimilarities and *pu*-spaces.

Proposition 2.6.1 *Let E be a G-semi-similarity, let us define* $|\ |_E : R \to [0, 1]$ *by setting*

$$|x|_E = 1 - E(x, x)$$

and $\delta_E : R \times R \to [0, 1]$ *by setting*

$$\delta_E(x, y)= \begin{cases} 0 & \text{if } E(x,y) = E(x,x) \wedge E(y,y) \\ 1 - E(x,y) & \text{if } E(x,y) < E(x,x) \wedge E(y,y) \end{cases}$$

for every $x, y \in R$. *Then* $(R, \delta_E, |\ |_E)$ *is a pu-space.*

Conversely, let $(R, \delta, |\ |)$ *be a pu-space and let us define* $E_{\delta, |\ |} : R \times R \to [0,1]$ *by setting*

$$E_{\delta, |\ |}(x, y) = 1 - (\delta(x, y) \vee |x| \vee |y|) \tag{2.7}$$

Then $E_{\delta, |\ |}$ *is a G-semi-similarity.*

This last proposition will be useful to describe the link of point-free geometry with fuzzy sets theory by a categorical point of view.

Before proceeding, let us summarize in a scheme all the structures we presented and all the connections between them we found.

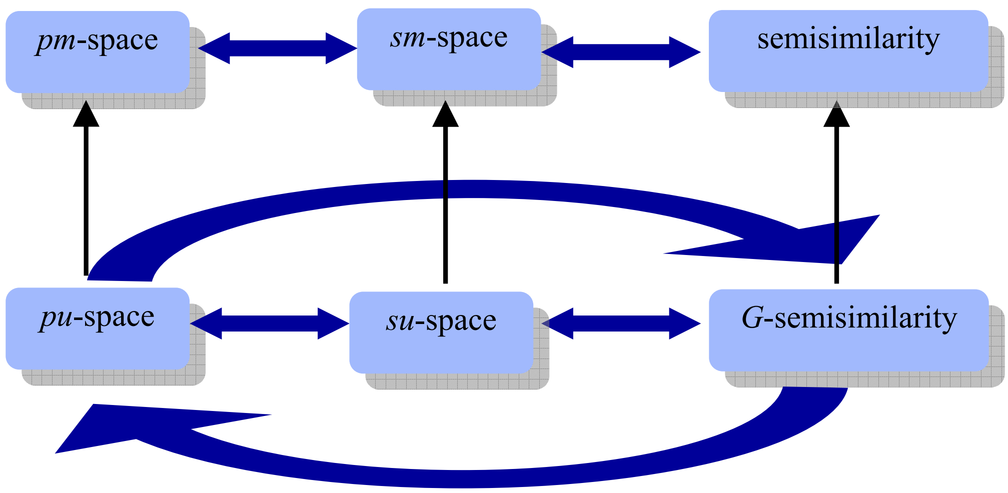


Figure 4. Scheme of the connections

In order to organize the class of semisimilarities into a category, we refer to the category *M*-SET* described by Hőhle in [28]. Namely, while Hőhle defines this category for any *GL*-monoid, we are interested only with the particular *GL*-monoid in [0,1] defined by the Gödel t-norm. In such a case we have the following simplified definition.

Definition 2.6.2 The *category of the G-semisimilarities* is the category ***GSS*** such that:

- the objects are structures (R, E) in which E is a G-semisimilarity;
- a morphism from (R, E) to (R', E') is a map $f : R \to R'$ satisfying the axioms

(M1) $E'(f(x), f(x)) \leq E(x, x)$

(M2) $E(x, y) \leq E'(f(x), f(y))$

for every $x,\ y \in R$.

Let us observe that from (M2) we have that

$$E(x, x) \leq E'(f(x), f(x))$$

and therefore, by (M1),

$$E(x, x) = E'(f(x), f(x)).$$

Now we define a category based on the class of *pu*-spaces.

Definition 2.6.3 The *category* ***PU*** of the *pu-spaces* is the category such that

- the objects are *pu*-spaces;
- a morphism from $(R, \delta, |\ |)$ to $(R', \delta', |\ |')$ is a map $f: R \rightarrow R'$ such that

(1) $\delta(x, y) \geq \delta'(f(x), f(y))$

(2) $|x| \geq |f(x)|'$

for every $x,\ y \in R$.

In both the categories the *composition* is the usual composition of maps and the *identities* are the identical maps.

Let us briefly observe that in the category ***PU***, morphisms preserve the definition of point we have given in Section 2.1. In fact, let us consider two abstraction processes $\langle x_n \rangle_{n\in\mathbb{N}}$, $\langle y_n \rangle_{n\in\mathbb{N}}$ and a morphism f in the category ***PU***. Then, if $\langle x_n \rangle \equiv \langle y_n \rangle \in [\langle x_n \rangle]$, it results that $lim_{\ n\rightarrow\infty}\delta(x_n, y_n) = 0$. Since $\delta(x_n, y_n) \geq \delta'(f(x_n), f(y_n))$ holds, then also $lim_{\ n\rightarrow\infty}\delta'(f(x_n), f(y_n)) = 0$. Therefore, $\langle f(x_n) \rangle \equiv \langle f(y_n) \rangle \in [\langle f(x_n) \rangle]$.

Proposition 2.6.1 enables us to associate any *G*-semisimilarity with a *pu*-space $(R, \delta_E, |\ |_E)$. This suggests the definition of a suitable functor from ***GSS*** to ***PU***.

Proposition 2.6.2 *Let us define the map F from* ***GSS*** *to* ***PU*** *by setting*

• $F((R, E)) = (R, \delta_E , |\ |_E)$

• $F(f) = f$.

*Then F is a functor from **GSS** to **PU**.*

Proof. We have only to prove that if f is a morphism from (R, E) to (R', E'), then f is a morphism from $(R, \delta_E, |\,|_E)$ to $(R', \delta_{E'}, |\,|_{E'})$. Indeed, it is immediate that

$$|f(x)|_{E'} = 1\text{-}E'(f(x), f(x)) = 1\text{-}E(x, x) = |x|_E.$$

To prove that

$$\delta_E(x, y) \geq \delta_{E'}(f(x), f(y)) \tag{2.8}$$

it is not restrictive to assume that $\delta_{E'}(f(x), f(y)) \neq 0$ and therefore that

$$E'(f(x), f(y)) < E'(f(x), f(x)) \wedge E'(f(y), f(y))$$

and

$$\delta_{E'}(f(x), f(y)) = 1\text{-}E'(f(x), f(y)).$$

In such a case, since

$$E(x, y) \leq E'(f(x), f(y)) < E'(f(x), f(x)) \wedge E'(f(y), f(y)) = E(x, x) \wedge E(y, y),$$

we have that $\delta_E(x, y) = 1\text{-}E(x, y)$. So, (2.8) is a trivial consequence of (M2).

□

Let us observe that in proving that F is a functor we obtain that

$$|f(x)|_{E'} = |x|_E\,. \tag{2.9}$$

On the other hand, it is easy to find a morphism h in ***PU*** such that $|f(x)|_{E'} < |x|_E$ for a suitable region. Then, the proposed functor is faithful, but not full. We can consider the subcategory ***PU**** of ***PU*** obtained by considering only the morphisms f satisfying (2.9). Proposition 2.6.1 suggests a definition of a functor from ***PU**** to ***GSS***.

Proposition 2.6.3 *Let us define the map F' from **PU*** to **GSS** by setting*

- $F'((R, \delta, |\,|)) = (R, E_\delta, |\,|)$
- $F(f) = f$.

*Then F' is a functor from **PU*** to **GSS**.*

Proof. Let $(R, \delta, |\ |)$ and $(R', \delta', |\ |')$ be two *pu*-spaces, let (R, E) and (R', E') be the structures, where the semi-similarities E and E' are defined by (2.7), and let f be a morphism from $(R, \delta, |\ |)$ to $(R', \delta', |\ |')$. Then

$$E'(f(x), f(x)) = 1-|f(x)|' = 1-|x| = E(x, x).$$

Moreover,

$$E(x, y) = 1-(\delta(x, y) \vee |x| \vee |y|) \leq 1-(\delta'(f(x), f(y)) \vee |f(x)|' \vee |f(y)|') = E'(f(x), f(y)).$$

□

2.7 A *G*-semisimilarity on the class of partial functions

Given two nonempty sets X and Y we denote by $F(X, Y)$ the class of partial functions from X to Y. If $f \in F(X, Y)$ we denote by D_f the domain of f and by U_f the complement of D_f, i.e. the set of elements in which f is not defined. Let f, g be elements of $F(X, Y)$, then the *equalizer* of f and g , is defined by

$$eq(f, g) = \{x \in X \,/\, x \in D_f \cap D_g\ , f(x)=g(x)\}.$$

The contrast between f and g is defined as the complement of the equalizer, i.e.

$$contr(f, g) = -\ eq(f, g).$$

Let us observe that

$$contr(f, g) = C_{fg} \cup U_f \cup U_g,$$

where

$$C_{fg} = \{x \in X \,/\, x \in D_f \cap D_g \ \ and \ \ f(x) \neq g(x)\}.$$

In other words, $contr(f, g)$ contains the elements on which f and g "*contrast*" and the elements in which either f or g is not defined. We can also interpret $contr(f, g)$ as the set of the elements in which either f and g actually contrast or they could contrast in successive extensions. In particular

$$contr(f, f) = U_f.$$

Definition 2.7.1 Let us consider a map *irl*: $X\to[0,1]$ we call *fuzzy subset of irrelevant elements*. Then the *irrelevancy degree of a set S*, is

$$Irl(S)= Inf\{irl(x)\ /\ x\in S\}.$$

We interpret $irl(x)$ as the "*degree of irrelevancy*" of an element x and $Irl(S)$ as a measure of the degree of validity of the claim "all the elements in S are irrelevant". Trivially, we have that for any pair S_1, S_2 of subsets of X,

$$Irl(S_1\cup S_2) = Irl(S_1) \wedge Irl(S_2).$$

Proposition 2.7.1 *Let $F(X, Y)$ be the class of partial functions and let E: $F(X, Y)\times F(X, Y)\to[0, 1]$ be defined by*

$$E(f, g) = Irl(contr(f, g)).$$

Then E is a G- semi-similarity.

Proof. (e1) is immediate. To prove (e2*), let us observe that for every f, g, $h\in X$,

$$C_{fg} \subseteq C_{fh}\cup C_{hg} \cup U_h$$

and therefore,

$$contr(f, g) \subseteq contr(f, h) \cup contr(h, g).$$

This entails

$$E(f, g) \geq E(f, h) \wedge E(h, g).$$

□

We interpret $E(f, g)$ as a measure of the truth degree of the claim "in all the relevant elements f and g are defined and coincide". Let us observe that

$$E(f, f) = Irl(U_f)$$

and therefore $E(f, f)$ is the valuation of the claim that f is defined in all the relevant elements. In particular, if f is total, then $E(f, f)$ is equal to 1, if f is totally undefined, i.e. $U_f = X$, then $E(f, f) = 0$.

Chapter 3
Approximate distances and incomplete information

Interval analysis is a basic tool to face the question of the errors arising either in measure processes or in approximate calculations, providing an upper and a lower bound for the exact solution of a problem (see, for example, [3] and [27]). In agreement with the ideas of interval analysis, in this chapter we introduce the concept of *approximate distance*, which extends the notion of distance taking into account the errors arising from the incomplete knowledge of the points. We do this by using interval-valued maps, which do not assign, to a pair of objects, a single value, but an interval representing the range in which the exact value lies. Moreover, developing Whitehead's ideas we introduce the approximate distance on spaces of regions, representing the incompleteness of the knowledge. Hence, we define an abstract structure of *interval semimetric space*, based on an interval-valued "distance".

Then, this construction is extended to the fuzzy setting, defining approximate distances between fuzzy sets. By means of interval sets, we also define approximate distances between rough sets. Finally an application of our interval approximation to a clustering procedure is given (see [6], [7]).

3.1 Preliminaries

We denote by $\mathbb{R}$ and $\mathbb{R}_0^+$, $I(\mathbb{R})$ and $I(\mathbb{R}_0^+)$ the set of real numbers, the set of non-negative real numbers, the set of closed intervals in $\mathbb{R}$, the set of closed intervals in $\mathbb{R}_0^+$, respectively. Also, we denote by $w([u, v])$, the *width* $v - u$ of a nonempty interval $[u, v]$ and by $\pi_1: I(\mathbb{R}_0^+)\to\mathbb{R}_0^+$ and $\pi_2: I(\mathbb{R}_0^+)\to\mathbb{R}_0^+$ the *first* and the *second projection* of an interval, respectively, i.e. the functions defined by $\pi_1([u, v]) = u$ and $\pi_2([u, v]) = v$.

In general it is possible to lift any operation on elements of a given set to an operation on its subsets (see, for example, [2] and [50]). In the same way, if $\perp$ is a total defined binary arithmetic operation in $\mathbb{R}$, then it is possible to extend this operation to the intervals (see [27], [34]):

$$[u, v] \perp [u', v'] = \{x \perp y \,/\, u \le x \le v \text{ and } u' \le y \le v'\},$$

for any $[u, v]$ and $[u', v']\in I(\mathbb{R}_0^+)$. In the case of the addition we simply have:

$$[u, v] + [u', v'] = [u + u', v + v'].$$

In the case of the difference, since $-[u', v'] = [-v', -u']$, we have

$$[u, v] - [u', v'] = [u - v', v - u'].$$

In the case of the product we have

$$[u, v]*[u', v']=[min(u * u', u * v', v * u', v * v'), max(u * u', u * v', v * u', v * v')].$$

Let us remark that if $0 \notin [u', v']$, then also the division $[u, v] / [u', v']$ is definable in an analogous way. In such a case we have

$$[u, v] / [u', v'] = [u, v]* [1/ u', 1/ v'].$$

Similar definitions are proposed for unary operations.

In addition to the usual inclusion relation $\subseteq$, we consider the following partial order:

$$[u, v] \le_I [u', v'] \text{ iff } u \le u' \text{ and } v \le v'. \tag{3.1}$$

Let us consider the map $e: u\in\mathbb{R} \to [u, u]\in I(\mathbb{R})$. Then e is an embedding of $(\mathbb{R}, +, *, 0, 1, \le)$ into $(I(\mathbb{R}),+, *, 0, 1, \le_I)$. In accordance, we call *interval extension* of $(\mathbb{R}, +, *, 0, 1, \le)$ the structure $(I(\mathbb{R}),+, *, 0, 1, \le_I)$. Hence, we will use u or the degenerate interval $[u, u]$, indistinctly. In particular, we write 0 and 1 to denote [0, 0] and [1, 1], respectively. Let us observe that, in dealing with intervals, some of the classical arithmetic rules fail. As an example, an interval has neither additive nor multiplicative inverse. Also, the distributive rule fails, whereas the operations $+$ and $*$ are associative and commutative and they have as neutral elements 0 and 1, respectively.

We can organize the set of real intervals as a metric space. Given two intervals $[u, v]$ and $[u', v'] \in I(\mathbb{R})$, a *distance* between them is defined as

$$d([u, v], [u', v']) = max\ \{|u - u'|, |v - v'|\}.$$

It is easy to show that d is a *metric* in $I(\mathbb{R})$. We can use this metric distance in order to introduce the notion of convergence of a sequence of intervals.

Definition 3.1.1 A sequence of intervals $\langle[u_n, v_n]\rangle_{n\in\mathbb{N}}$ in $I(\mathbb{R})$ *converges* to $[u, v]$ if $lim_{n\to\infty}\, d([u_n, v_n], [u, v])=0$.

From the definition of the distance d, it follows that $\langle[u_n, v_n]\rangle_{n\in\mathbb{N}}$ converges to $[u, v]$ if $lim_{n\to\infty}\, u_n = u$ and if $lim_{n\to\infty}\, v_n = v$, i.e. the end points of $\langle[u_n, v_n]\rangle_{n\in\mathbb{N}}$ converge to the end points of the limit interval. Particularly, if $\langle[u_n, v_n]\rangle_{n\in\mathbb{N}}$ is a decreasing sequence of intervals with respect to the inclusion, then $\langle[u_n, v_n]\rangle_{n\in\mathbb{N}}$ converges to $\cap_{n\in\mathbb{N}}\, [u_n, v_n] = [Sup_{\,n\in\mathbb{N}}\, u_n, Inf_{\,n\in\mathbb{N}}\, v_n]$. If $lim_{n\to\infty}\, w([u_n, v_n]) = 0$, then the sequence $\langle[u_n, v_n]\rangle_{n\in\mathbb{N}}$ converges to a degenerate interval $[l, l]$ and so we write $lim_{n\to\infty}\, ([u_n, v_n]) = l$.

3.2 Interval semimetric spaces

We define an abstract structure on a partially ordered set $(R, \leq_R)$ by introducing a non-negative, interval-valued map Δ on $R\times R$. This structure, defined by a set of axioms, results a generalization, in terms of intervals, of pseudometric spaces. The intended interpretation is that the elements in R are regions in a geometrical space, $\leq_R$ is the inclusion relation and that Δ is an "approximate" distance.

Definition 3.2.1 An *interval semimetric space*, briefly ***ISM**-space*, is a structure $(R, \leq_R, \Delta)$, where $(R, \leq_R)$ is a partially ordered set, $\Delta: R\times R \to I(\mathbb{R}_0^+)$ is an interval-valued map, such that, for every $x, y, z \in R$, the following axioms hold:

A1) $\Delta(x, x) * [0, 1] = \Delta(x, x)$,

A2) $\Delta(x, y) = \Delta(y, x)$;

A3) $\Delta(x, y) - \Delta(z, z) \leq_I \Delta(x, z) + \Delta(z, y)$;

A4) $\Delta(x, y) - \Delta(x, y) \leq_I \Delta(x, x) + \Delta(y, y)$;

A5) $x \leq_R x', y \leq_R y' \Rightarrow \Delta(x, y) \subseteq \Delta(x', y')$.

We call *regions* the elements in R and we call *weight function* the map $p: R \to \mathbb{R}_0^+$ defined by

$$p(x) = \pi_2(\Delta(x, x)),$$

for every $x \in R$.

Also, we can interpret an element in R as an "incomplete piece of information" and $\Delta(x, y)$ as an approximate measure of how two pieces of information x and y are close, i.e. as the available knowledge about the actual but unknown distance between them. Then, we can interpret weight $p(x)$ as a measure of the degree of information carried on by a region, that is, in other words, a measure of the completeness of x.

From axioms A1), A4) and A5) we can derive the following properties of p:

- $\Delta(x, x) = [0, p(x)]$; (3.2)

- $w(\Delta(x, y)) \leq p(x) + p(y)$; (3.3)
- $x \leq_R y \Rightarrow p(x) \leq p(y)$. (3.4)

By setting $x \leq_R x$ and $x \leq_R y$ in A5), it results that

- $x \leq_R y \Rightarrow \pi_1(\Delta(x, y)) = 0$.

Let us briefly explain the meaning of the axioms we propose. Axiom A1) is a sort of weakening of reflexivity for the distance Δ, taking in account the completeness of the information x. More precisely, it states that the information we have about a region x is between 0 and $p(x)$, as we can also see in the deriving property (3.2). We have the reflexivity only in the case of complete information. Axiom A2) gives the symmetry for Δ. Axiom A3) represents a generalized triangular inequality, in which we consider also the measure of the completeness of the middle region z. Axiom A4) can be interpreted as a sort of constraint for the width of the interval $\Delta(x, y)$, as we can realize also from the deriving property (3.3). In other words, the more complete the information about the regions, the less approximate the distance between them. Finally, axiom A5) indicates the order-preserving characteristic of Δ, with respect to the inclusion relation between intervals.

Definition 3.2.2 Two ***ISM***-spaces $(R, \leq_R, \Delta)$ and $(R', \leq_{R'}, \Delta')$ are *isomorphic* if there exists a bijective map $\varphi: R \to R'$ which preserves the structure, i.e., for every $x, y \in R$:

1) $x \leq_R y \Rightarrow \varphi(x) \leq_{R'} \varphi(y)$

2) $\Delta'(\varphi(x), \varphi(y)) = \Delta(x, y)$

If $(R, \leq_R, \Delta)$ and $(R', \leq_{R'}, \Delta')$ are isomorphic, then $p'(\varphi(x)) = p(x)$, where p and p' are the corresponding weight functions.

Definition 3.2.3 An *interval metric space*, briefly ***IM***-space, is an interval semimetric space which satisfies the axiom:

A6) $\Delta(x, y) = 0 \Leftrightarrow x = y$ and x and y are atoms,

where we call *atoms* the minimal elements x of R such that $p(x) = 0$.

Let us observe that, as in the case of metric spaces, any subset of an ***ISM***-space is an ***ISM***-space, too. ***ISM***-spaces (***IM***-spaces) can be regarded as a generalization of pseudometric spaces (metric spaces).

Proposition 3.2.1 *Pseudometric spaces* (*metric spaces*) *coincide with the* ***ISM****-spaces* (*IM-spaces*) *in which all the regions are atoms, i.e. the order relation is the identity and all the weights are equal to zero.*

Now we enunciate a Theorem in order to emphasize that the notion of ***ISM***-space can be presented in two different but equivalent ways: either by two real valued-maps d and D or by an interval-valued map Δ.

Theorem 3.2.1 *Let* $(R, \leq_R, \Delta)$ *be an* ***ISM****-space and let d and D be defined, for any* $x, y \in R$, *by:*

$$d(x, y) = \pi_1(\Delta(x, y)) \text{ and } D(x, y) = \pi_2(\Delta(x, y)).$$

Then, for $x, y \in R$, $d(x, y) \leq D(x, y)$, *d and D are order-reversing and order-preserving, respectively. Moreover, for* $x, y, z \in R$, *the following properties hold*

$P1)$ $d(x, x) = 0$,

$P2a)$ $d(x, y) = d(y, x)$,

$P2b)$ $D(x, y) = D(y, x)$,

$P3)$ $d(x, y) \leq d(x, z) + d(z, y) + D(z, z)$,

$P4)$ $D(x, y) \leq D(x, z) + D(z, y)$,

$P5)$ $0 \leq D(x, y) - d(x, y) \leq D(x, x) + D(y, y)$.

Conversely, let d and D be two maps from $R \times R$ *to* $\mathbb{R}_0^+$ *order-reversing and order-preserving, respectively, such that* $d(x, y) \le D(x, y)$, *for any* $x, y \in R$, *and verifying* $P1) - P5)$. *Let* Δ *be defined, for any* $x, y \in R$, *by:*

$$\Delta(x, y) = [d(x, y), D(x, y)].$$

Then $(R, \le_R, \Delta)$ *is an* ***ISM****-space.*

Proof. Properties $P1)$, $P2a)$ and $P2b)$ follow from A1) jointly with A2), trivially. Moreover, by definition of d and D, A3) becomes

$$[d(x, y), D(x, y)] - [d(z, z), D(z, z)] \le_I [d(x, z), D(x, z)] + [d(z, y), D(z, y)], \text{ i.e.}$$

$$[d(x, y) - D(z, z), D(x, y) - d(z, z)] \le_I [d(x, z) + d(z, y), D(x, z) + D(z, y)].$$

Therefore properties $P3)$ and $P4)$ can be easily deduced applying property $P1)$ and the definition of $\le_I$, given in (3.1). Similarly, axiom A4) becomes

$$[d(x, y), D(x, y)] - [d(x, y), D(x, y)] \le_I [0, D(x, x)] + [0, D(y, y)], \text{ i.e.,}$$

$$[d(x, y) - D(x, y), D(x, y) - d(x, y)] \le_I [0, D(x, x) + D(y, y)].$$

Therefore, again by the definition of $\le_I$, we obtain $P5)$. Finally, axiom A5) ensures d and D to be order-reversing and order-preserving, respectively.

To prove the second part of the theorem, let us observe that axioms A1) - A4) follow from properties $P1) - P5)$, trivially. Axiom A5) can be easily deduced applying the definition of Δ and the attribute of d and D to be order-reversing and order-preserving, respectively.

□

3.3 Abstraction processes in *ISM*-spaces

In point-free ***ISM***-spaces, i.e. without atoms, a suitable definition of abstraction process enables us to define points and a distance between them (see [5]). Thus, a metric space can be associated to any interval semimetric space.

Definition 3.3.1 Let $(R, \leq_R, \Delta)$ be an ***ISM***-space. An *abstraction process* is a sequence $\langle x_n \rangle_{n\in\mathbb{N}}$ in $(R, \leq_R, \Delta)$ such that

a) $\langle x_n \rangle_{n\in\mathbb{N}}$ is order-reversing;

b) $lim_{n\to\infty} p(x_n) = 0$.

For such a $\langle x_n \rangle_{n\in\mathbb{N}}$ by this definition, by the property (3.2) of p and by axiom A5), we also have the following:

c) $\forall \varepsilon > 0, \exists \nu \,/\, \forall n, m > \nu, \pi_2(\Delta(x_n, x_m)) < \varepsilon$ (or, analogously, $\Delta(x_n, x_m) \subseteq [0, \varepsilon]$).

We denote by $AP(R)$ the class of abstraction processes of $(R, \leq_R, \Delta)$.

In order to define a pseudometric in $AP(R)$, we need to require the following axiom:

A7) $AP(R) \neq \emptyset$.

Proposition 3.3.1 *Let $(R, \leq_R, \Delta)$ be an **ISM**-space such that A7) holds and for any pair $\langle x_n \rangle_{n\in\mathbb{N}}$, $\langle y_n \rangle_{n\in\mathbb{N}}$ of abstraction processes, let δ: $AP(R)\times AP(R) \to \mathbb{R}^+$ be a map defined by*

$$\delta(\langle x_n \rangle_{n\in\mathbb{N}}, \langle y_n \rangle_{n\in\mathbb{N}}) = lim_{n\to\infty}\Delta(x_n, y_n).$$

Then $(AP(R), \delta)$ is a pseudometric space.

Proof. Let us observe that, if $\langle x_n \rangle_{n\in\mathbb{N}}$ and $\langle y_n \rangle_{n\in\mathbb{N}}$ are abstraction processes, then, by A5), $\langle \Delta(x_n, y_n) \rangle_{n\in\mathbb{N}}$ is a decreasing sequence of intervals, whose limit is $\cap_{n\in\mathbb{N}} \Delta(x_n, y_n)$; besides, by (3.3), when $n\to\infty$, the width $w(\Delta(x_n, y_n)) \to 0$ and so the sequence $\langle \Delta(x_n, y_n) \rangle_{n\in\mathbb{N}}$ is convergent to a real number. Let $\langle x_n \rangle_{n\in\mathbb{N}}$ be an element of $AP(R)$. By definition of δ, by axiom A1) and by b) of the Definition 3.3.1, we have $\delta(\langle x_n \rangle_{n\in\mathbb{N}}, \langle x_n \rangle_{n\in\mathbb{N}}) = lim_{n\to\infty}\Delta(x_n, x_n) = lim_{n\to\infty}([0, p(x_n)] = 0$. So, the reflexivity for δ

is proved. The symmetry is trivial by axiom A2). Finally, to prove the triangular inequality, let us observe that, if $\langle x_n\rangle_{n\in\mathbb{N}}$, $\langle y_n\rangle_{n\in\mathbb{N}}$ and $\langle z_n\rangle_{n\in\mathbb{N}}$ are elements of $AP(R)$, then, by axiom A3) and by b) of the Definition 3.3.1, we have

$$\delta(\langle x_n\rangle_{n\in\mathbb{N}}, \langle y_n\rangle_{n\in\mathbb{N}}) = lim_{n\to\infty}\Delta(x_n, y_n) \leq_I lim_{n\to\infty}(\Delta(x_n, z_n) + \Delta(z_n, y_n) + p(z_n))$$

$$=lim_{n\to\infty}\Delta(x_n, z_n) + lim_{n\to\infty}\Delta(z_n, y_n) + lim_{n\to\infty}p(z_n)$$

$$=\delta(\langle x_n\rangle_{n\in\mathbb{N}}, \langle z_n\rangle_{n\in\mathbb{N}}) + \delta(\langle z_n\rangle_{n\in\mathbb{N}}, \langle y_n\rangle_{n\in\mathbb{N}})$$

□

Definition 3.3.2 Given an interval semimetric space $(R, \leq_R, \Delta)$, we call *metric space associated to* $(R, \leq_R, \Delta)$ the metric space (M, δ') obtained as a quotient of $(AP(R), \delta)$ modulo the relation $\equiv$ defined by setting

$$\langle x_n\rangle_{n\in\mathbb{N}} \equiv \langle y_n\rangle_{n\in\mathbb{N}} \text{ iff } \delta(\langle x_n\rangle_{n\in\mathbb{N}}, \langle y_n\rangle_{n\in\mathbb{N}})= 0.$$

We call *points* the elements of M, i.e. the equivalence classes

$$[\langle x_n\rangle_{n\in\mathbb{N}}] = \{\langle y_n\rangle_{n\in\mathbb{N}} \in AP(R) \,/\, \langle x_n\rangle_{n\in\mathbb{N}} \equiv \langle y_n\rangle_{n\in\mathbb{N}}\}.$$

As a consequence, we can define a distance between two points P and Q by setting

$$\delta'(P, Q) = \delta(\langle x_n\rangle_{n\in\mathbb{N}}, \langle y_n\rangle_{n\in\mathbb{N}}) = lim_{n\to\infty}\Delta(x_n, y_n),$$

where $\langle x_n\rangle_{n\in\mathbb{N}}$, $\langle y_n\rangle_{n\in\mathbb{N}}$, belonging to $AP(R)$, are two representatives of P and Q, respectively. In this way, it is always possible to associate any interval semimetric space with a metric space.

Let us note that every atom defines a point with respect to the definition above.

3.4 Canonical models of *ISM*-spaces

Now we yield models of ***ISM***-spaces by using the canonical lower and upper distances between subsets of a pseudometric space (M, δ). Really, to search for a suitable system of axioms of interval-valued distances, we referred just to these particular examples.

Let C be a nonempty class of bounded, nonempty subsets of M. The canonical *infimum* and *supremum* distances d and D are defined, respectively, by setting

$$d(x, y) = inf\{\delta(X, Y) : X \in x, Y \in y\} \tag{3.5}$$

and

$$D(x, y) = sup\{\delta(X, Y) : X \in x, Y \in y\} \tag{3.6}$$

for every $x, y \in C$. Besides we call *diameter*

$$|x| = sup\{\delta(X, Y) : X, Y \in x\} = D(x, x).$$

Proposition 3.4.1 *Given a nonempty class C of bounded, nonempty subsets of a pseudometric space (M, δ), let d and D be the infimum and supremum distances and let $\subseteq$ be the usual inclusion relation between subsets. Let Δ: $C \times C \to I(\mathbb{R}_0^+)$ be defined by*

$$\Delta(x, y) = [d(x, y), D(x, y)].$$

*Then $(C, \subseteq, \Delta)$ is an **ISM**-space, whose weight function coincides with the diameter.*

Proof. By the equivalence shown in Theorem 3.2.1, we have to prove that d and D satisfy $P1)$ - $P5)$. Now, $P1)$, $P2a)$ and $P2b)$ are immediate. To prove $P3)$, let x, y and z be subsets of M, $X \in x$, $Y \in y$, Z and $Z' \in z$; then

$d(x, y) \le \delta(X, Y) \le \delta(X, Z) + \delta(Z, Z') + \delta(Z', Y) \le \delta(X, Z) + \delta(Z', Y) + |z|$. Consequently,

$$d(x, y) \le d(x, z) + d(z, y) + |z|.$$

To prove $P4)$, let x, y and z be subsets of M, $X \in x$, $Y \in y$, $Z \in z$; then

$$\delta(X, Y) \le \delta(X, Z) + \delta(Z, Y) \le D(x, z) + D(z, y);$$

hence

$$D(x, y) \le D(x, z) + D(z, y).$$

To prove $P5)$, let x and y be subsets of M, $X, X' \in x$, $Y, Y' \in y$; then

$\delta(X, Y) \le \delta(X, Y') + \delta(Y', Y) \le \delta(Y', X') + \delta(X', X) + \delta(Y', Y) \le \delta(Y', X') + |x| + |y|$.

It follows that

$$\delta(X, Y) \leq inf\{\delta(X, Y) : X \in x, Y \in y\}+|x| +|y| = d(x, y)+|x| +|y|.$$

Further, it follows that

$$D(x, y) = sup\{\delta(X, Y) : X \in x, Y \in y\}\leq d(x, y)+|x| +|y|.$$

Besides, by definition, $d(X, Y) \leq D(x, y)$. So

$$0 \leq D(x, y) - d(x, y) \leq |x| +|y|.$$

□

Let us observe that if C is closed with respect to the union, then there are two further interesting properties verified by d and D, a sort of *distributive* laws:

$$d(x_1\cup x_2, y_1\cup y_2) = d(x_1, y_1) \wedge d(x_1, y_2) \wedge d(x_2, y_1) \wedge d(x_2, y_2) \quad (3.7)$$

and

$$D(x_1\cup x_2, y_1\cup y_2) = D(x_1, y_1) \vee D(x_1, y_2) \vee D(x_2, y_1) \vee D(x_2, y_2). \quad (3.8)$$

If we interpret two sets x and y in C as constraints on two unknown points X and Y, then the ”approximate” distance $\Delta(x, y)$ yields a constraint on the possible distance between X and Y.

Definition 3.4.2 The structure $(C, \subseteq, \Delta)$ is said to be a *canonical* ***ISM***-space.

It is possible to construct several canonical spaces, depending on the choice of C. Particularly, we are interested in the class C of all closed, bounded, regular, nonempty subsets of a pseudometric space (M, δ). We make this choice because, in according to Whitehead's works (see [46], [47], [48]), we exclude the ”abstract” geometrical entities like points, lines that can be obtained by an abstraction process.

Henceforth, in this section and in the following ones, we denote by C such a class, even though some of the following results can be obtained in larger classes.

Let us observe that one of the most well-known distances between sets, the Hausdorff distance (see (1.6)), falls just into the interval whose end points are the infimum and supremum distances, as we are going to show in the next proposition.

Proposition 3.4.2*Let C be a nonempty class of bounded, nonempty subsets of a be a pseudometric space* (M, δ) *and let* δ_H *be the related Hausdorff distance. Then* $d(x, y) \leq \delta_H(x, y) \leq D(x, y)$, *for every* $x, y \in C$.

Proof. By definition of δ_H, we have
$\delta_H(x, y) \geq sup_{X\in x}\, \delta(X, y) \geq \delta(X, y)$ for every $X \in x$, i.e.
$\delta_H(x, y) \geq inf_{Y\in y} \delta(X, Y)$ for every $X \in x$. Hence,

$$\delta_H(x, y) \geq inf\{\delta(X, Y) \,/\, X\in x,\, Y\in y\} = d(x, y).$$

On the other hand,
$D(x, y) \geq \delta(X, Y)$, for every $X \in x$, and $Y \in y$. Therefore,
$D(x, y) \geq inf_{Y\in y}\delta(X, Y)$ for every $X \in x$, i.e. $D(x, y) \geq \delta(X, y)$, for every $X \in x$. Particularly $D(x, y) \geq sup_{X\in x}\delta(X, y) = e_\delta(x, y)$. Analogously this inequality holds for $e_\delta(y, x)$, so

$$D(x, y) \geq \delta_H(x, y).$$

□

Now we want to show that, starting form a complete pseudometric space, (M, δ), the resulting metric space associated to $(C, \subseteq, \Delta)$ is isometric to (M, δ), recalling that we call two metric spaces *isometric* if there exists a bijective isometry between them.

So, let us denote by (M_C, δ') the metric space associated to the canonical model $(C, \subseteq, \Delta)$.

Trivially, if (M, δ) is a metric space, then for every isolated point $P \in M$, the sequence constantly equal to P is an abstraction process. Particularly, if δ is the discrete metric, then this happens for every $P \in M$.

Given $P \in M$ and $n \in \mathbb{N}$, let us denote by $B_n(P)$ the closure of the open ball centered in P, with radius $1/n$. Let us observe that $\langle B_n(P)\rangle_{n\in\mathbb{N}}$ is an order-reversing sequence of regular sets such that $lim_{n\to\infty}|B_n(P)| = lim_{n\to\infty}\langle 2/n\rangle = 0$. Therefore $\langle B_n(P)\rangle_{n\in\mathbb{N}}$ is an abstraction process and so it also results that $\forall\varepsilon>0$, $\exists\, \nu$ / $\forall n, m>\nu$, $\pi_2(\Delta(B_n(P), B_m(P))) = D(B_n(P), B_m(P))) <\varepsilon$ (or, equivalently, $\Delta(B_n(P), B_m(P)) \subseteq [0, \varepsilon]$).

We can define the map i: $M \to M_C$ such that

$$i(P) = [\langle B_n(P)\rangle_{n\in\mathbb{N}}]. \qquad (3.9)$$

Proposition 3.4.3 *Let* (M, δ) *be a pseudometric space and let* (M_C, δ') *be the metric space associated to the canonical* ***ISM****-space* $(C, \subseteq, \Delta)$*. Let i be the map defined as in* (3.9)*. Then*

i) $\delta(P, Q) = \delta'(i(P), i(Q))$*, for every* $P, Q \in M$*, i.e. i is an isometry;*

ii) *if* (M, δ) *is a metric space, then i is an injection;*

iii) *if* (M, δ) *is a complete metric space, then the isometry i is a bijection, i.e.* (M, δ) *and* (M_C, δ') *are isometric.*

Proof. To prove ***i)***, let $P, Q \in M$, hence $[\langle B_n(P)\rangle_{n\in\mathbb{N}}] = i(P)$ and $[\langle B_n(Q)\rangle_{n\in\mathbb{N}}] = i(Q)$; then we have

$$\delta'(i(P), i(Q)) = lim_{n\to\infty}\Delta(B_n(P), B_n(Q))$$
$$=lim_{n\to\infty}[d(B_n(P), B_n(Q)), D(B_n(P), B_n(Q))] = \delta(P, Q).$$

Indeed, since the two sequences of end points $\langle d(B_n(P), B_n(Q))\rangle_{n\in\mathbb{N}}$ and $\langle D(B_n(P), B_n(Q))\rangle_{n\in\mathbb{N}}$ both converge to the real number $\delta(P, Q)$, by Definition 3.1.1, the limit interval is just $\delta(P, Q)$.
To prove ***ii*)**, let us recall that an isometry between a metric space and a pseudometric space is always injective. So, the map i results injective, trivially. Finally, in order to show that i is surjective, let us take a point Q $\in$ (M_C, δ') represented by the sequence $\langle x_n\rangle_{n\in\mathbb{N}}$ and let us fix $P_n \in x_n$ for each $n\in\mathbb{N}$. Since evidently $\delta(P_n, P_m) \leq D(x_n, x_m)$ for each n, $m\in\mathbb{N}$ and $[\langle x_n\rangle_{n\in\mathbb{N}}]$ is an abstraction process, then $\langle P_n\rangle_{n\in\mathbb{N}}$, where $P_n \in x_n$ for each $n\in\mathbb{N}$, is a δ-Cauchy sequence in M. So the hypothesis of completeness for (M, δ) implies that $\langle P_n\rangle_{n\in\mathbb{N}}$ converges to a point $P \in M$. Now $i(P) = Q$. Indeed, considered the ball $B_n(P)$ for each $n\in\mathbb{N}$, if we observe that $d(x_n, y_n) \leq \delta(P_n, P)$ we have $lim_{n\to\infty}\Delta(x_n, y_n) = lim_{n\to\infty}d(x_n, y_n) = 0$.

□

We concude this section by observing that the construction of an ***ISM***-space, starting from a metric space, suggests to organize the class of ***ISM***-spaces into a category and to find a link between this one and a category of metric spaces.

We can consider the category ***M*** in which the objects are metric spaces (M, δ) and the morphisms are bijective isometries between metric spaces. Besides, we can consider the category ***ISM*** in which the objects are ***ISM***-spaces and the morphisms are isomorphisms between these structures. In both the categories the composition is the usual composition of maps and the identities are the identical maps.

Proposition 3.4.4 *Let i:* $(M_1, \delta_1) \rightarrow (M_2, \delta_2)$ *be a morphism in the category* ***M***. *Let F be a map from* **M** *to* ***ISM***, *defined by setting*

- $F((M, \delta)) = (C, \subseteq, \Delta)$ *obtained as in Proposition 3.4.1, for every object* (M, δ) *in* ***M***;

- $F(i)$: $x \in F((M_1, \delta_1)) \rightarrow i(x) \in F((M_2, \delta_2))$, *for every morphism i in M.*

Then, F is a functor from ***M*** *to* ***ISM****.*

Proof. By Proposition 3.4.1, the image $(C, \subseteq, \Delta)$ via F of a metric space (M, δ) is an ***ISM***-space; so it is an object of the category M. The map $F(i)$ results to be an isomorphism between ***ISM***-spaces, so it is a morphism in the category ***ISM***. Trivially, F preserves the identity and the composition between the morphisms.

□

3.5 Approximate distances between fuzzy sets

In this section we define approximate distances between fuzzy sets by means of closed *λ-cuts* of fuzzy subsets and of hypographs.

First, let us observe that the notion of closed *λ-cut* gives a general way to extend maps from sets into real numbers, for instance distances between sets, measures, diameters and so on. This procedure is different from the one based on Zadeh's extension principle, as we will see.

Definition 3.5.1 Let $S_1,\ldots, S_n$ be nonempty sets and let $D_f \subset P(S_1)\times,\ldots,\times P(S_n)$ be the domain of a real-valued monotone map $f : D_f \rightarrow \mathbb{R}$. Then the *canonical extension* of f is the map $f^*: D^*_f \rightarrow \mathbb{R}$, where

$D^*_f=\{(s_1,\ldots,s_n)\in F(S_1)\times\ldots\times F(S_n) \,/ (C(s_1,\lambda),\ldots, C(s_n, \lambda)) \in D_f$ for any $\lambda\in [0,1]\}$

and

$$f^*(s_1,\ldots,s_n) = \int_0^1 f(C(s_1, \lambda),\ldots, C(s_n, \lambda))d\lambda.$$

Let us observe that the function $f(C(s_1, \lambda),\ldots, C(s_n, \lambda))$ is monotone, then the Lebesgue integral, we are considering, always exists. Moreover f^* extends f in the sense that

$$f(X_1,\ldots, X_n) = f^*(\chi_{X_1},\ldots,\chi_{Xn}),$$

for every $X_1 \in P(S_1),\ldots, X_n \in P(S_n)$.

Particularly, let (M, δ) be a metric space and let d and D be the canonical infimum and supremum distances, defined in (3.5) and in (3.6).

Then, their extensions $d^* : C^*\times C^*\to \mathbb{R}$ and $D^*:C^*\times C^*\to \mathbb{R}$ are defined, respectively, by

$$d_p^*(s, s')= (\int_0^1 d\,(C(s,\lambda), C(s', \lambda))^p d\lambda)^{1/p}$$

and

$$D_p^*(s, s')= (\int_0^1 D\,(C(s,\lambda), C(s', \lambda))^p d\lambda)^{1/p},$$

for every $s, s' \in C^*$, where $C^* = \{s \in F(M)\colon C(s,\lambda) \in C \text{ for any } \lambda\in [0,1]\}$.

The extension of the diameter function is defined by

$$|s|_p^* = (\int_0^1 p((C(s,\lambda))^p\, d\lambda)^{1/p}.$$

Proposition 3.5.1 *Let Δ_p^*: $C^*\times C^*\to I(\mathbb{R}_0^+)$ be a map defined by*

$$\Delta_p^*(s, s') = [d_p^*(s, s'), D_p^*(s, s')].$$

*Then $(C^*, \subseteq, \Delta_p^*)$ is an **ISM**-space which extends the canonical **ISM**-space $(C, \subseteq, \Delta)$ and whose weight is the canonical extension of the diameter function.*

Proof. To prove that $(C^*, \subseteq, \Delta_p^*)$ is an ***ISM***-space, it suffices to observe that properties of d and D are extended to d_p^* and D_p^* by properties of integrals. Successively, to prove that this space is an extension of the canonical one, it is enough to consider the embedding e: $C \to C^*$ defined by $e(x) = \chi_X$, for every $x \in C$, where χ_X is the characteristic function related to x. Besides,

- $\chi_X \le \chi_y \Leftrightarrow x \subseteq y$;
- $\Delta_p^*(\chi_X, \chi_y)=[d_p^*(\chi_X, \chi_y), D_p^*(\chi_X, \chi_y)]=[d(x, y), D(x, y)]=\Delta(x, y)$.

□

Let us point out that in literature we find another distance based on λ-cuts, the *d_p-metric*. It is defined on the class of normal, convex, upper-semicontinous fuzzy sets having support with compact closure, as (see[11]):

$$d_p(s, s') = \left(\int_0^1 \delta_H(C(s,\lambda), C(s', \lambda))^p d\lambda\right)^{1/p},$$

for every $1 \le p < \infty$ and

$$d_\infty(s, s') = sup_{0\le\lambda\le1} \delta_H(C(s,\lambda), C(s', \lambda)),$$

where δ_H is the Hausdorff metric. Let us observe that, as an obvious consequence of Proposition 3.4.2, the following inequality holds:

$$d_p^*(s, s') \le d_p(s, s') \le D_p^*(s, s').$$

This is not the only way of operating. In literature we can find different methods to extend non-fuzzy mathematical concepts to a fuzzy setting, such as the methods based on the Zadeh's extension principle, which is one of the most basic ideas of fuzzy set theory. According to it, we find other distances between fuzzy sets, such as the fuzzy distance, defined as (see [12]):

$$\overline{d}^*_{(s,s')}(r) = sup_{\delta(P,\,Q)=r}\ min(s(P), s'(Q)),$$

for every $r \in \mathbb{R}^+$ and for every s, s' fuzzy sets on M. The distance $\overline{d}^*$ is a mapping from $\Im(M)\times\Im(M)$ to the set of fuzzy sets on $\mathbb{R}^+$ and it can be viewed as a fuzzy measure of dissimilarity between fuzzy sets. For connected (classical) subsets of M, this distance, as Δ_1^* (obtained from Δ_p^*, with p=1), becomes an ordinary interval, whose end points are the shortest and greatest distance between the subsets, respectively, and then it coincides with the canonical approximate distance Δ.

Now, we can construct, as usual, the metric space of points associated to the ***ISM***-space we have defined. In order to do this, let us shrink to the case of $p=1$, which can be generalized. So we consider the ***ISM***-space $(C^*, \subseteq, \Delta_1^*)$. Let us first observe that the set of the abstraction processes of C^*, that we denote by $AP(C^*)$, is nonempty. In fact, the sequence $\langle \chi_{B_n(P_0)} \rangle$, where $\chi_{B_n(P_0)}$ is the characteristic function related to the closure of the open ball in M with radius $1/n$ and center P_0, is an abstraction process. Precisely, we have that $lim_{n\to\infty}\, p_1^*(\chi_{B_n(P_0)}) = lim_{n\to\infty}\, D_1^*(\chi_{B_n(P_0)}, \chi_{B_n(P_0)}) = lim_{n\to\infty}\, D(B_n(P_0), B_n(P_0)) = lim_{n\to\infty}|B_n(P_0)| = lim_{n\to\infty} 2/n = 0$; besides $\langle \chi_{B_n(P_0)} \rangle$ is a decreasing sequence. Thus, we can define the pseudometric distance

$$\delta^*(\langle s_n\rangle, \langle s'_n\rangle) = lim_{n\to\infty}\Delta_1^*(s_n, s'_n),$$

for every s_n, $s'_n \in C^*$, and the associated metric space, denoted by (M_{C^*}, δ'^*).

Proposition 3.5.2 *Let* (M, δ) *be a metric space. The metric spaces* (M_{C}, δ') *and* (M_{C^*}, δ'^*)*, associated to* $(C^*, \subseteq, \Delta_1^*)$ *and* $(C, \subseteq, \Delta)$*, respectively, are isometric.*

Proof. Let us define the map $\varphi : (M_C, \delta') \to (M_{C^*}, \delta'^*)$ by

$$\varphi([\langle x_n\rangle_{n\in\mathbb{N}}]) = [\langle \chi_{x_n} \rangle_{n\in\mathbb{N}}].$$

It's a routine to show that φ preserves distances. In fact, it suffices to observe that $\delta'^*(\langle \chi_{x_n} \rangle_{n\in\mathbb{N}}, \langle \chi_{y_n} \rangle_{n\in\mathbb{N}}) = lim_{n\to\infty}\Delta_1^*(\chi_{x_n}, \chi_{y_n}) = lim_{n\to\infty}\Delta(x_n, y_n) = \delta(\langle x_n\rangle_{n\in\mathbb{N}}, \langle y_n\rangle_{n\in\mathbb{N}})$.

Moreover, let $Q^* = [\langle s_n\rangle_{n\in\mathbb{N}}] \in M_{C^*}$, where $\langle s_n\rangle_{n\in\mathbb{N}}$ is a decreasing sequence of fuzzy normal subsets such that $lim_{n\to\infty}\, p^*(s_n) = 0$.

Since $\langle s_n \rangle_{n \in \mathbb{N}}$ is decreasing, there exists a point P in M such that $s_n(P) = 1$, $\forall n \in \mathbb{N}$. Let $B_n(P)$ be the closed ball in M centered in P with radius $1/n$. The image $\varphi([\langle B_n(P) \rangle_{n \in \mathbb{N}}]) = [\langle s_n \rangle_{n \in \mathbb{N}}]$. Equivalently $lim_{n \to \infty} \Delta_1{}^*(\chi_{B_n(P)}, s_n) = 0$. Let us observe that $d_1{}^*(\chi_{B_n(P)}, s_n) = 0$, $\forall n \in \mathbb{N}$, and, by property $|D_1{}^*(s, s') - d_1{}^*(s, s')| \leq |\, s| + |\, s'|$, it follows $D_1{}^*(\chi_{B_n(P)}, s_n) \leq |\, \chi_{B_n(P)}| + |\, s_n|$ and therefore $lim_{n \to \infty} D_1{}^*(\chi_{B_n(P)}, s_n) = 0$.
The function φ is injective because it is an isometry defined on a metric space.

□

Corollary 3.5.1 *If* (M, δ) *is a complete metric space, then* (M_{C^*}, δ'^*) *is isometric to* (M, δ).

Proof. If (M, δ) is a complete metric space, then the diagram

$$\begin{tikzcd} & (M, \delta) \arrow[dl, "i"'] \arrow[dr, "\varphi \circ i"] & \\ (M_C, \delta') \arrow[rr, "\varphi"] & & (M_{C^*}, \delta'^*) \end{tikzcd}$$

is commutative.

□

A different definition of approximate distance between fuzzy subsets can be obtained by the notion of *hypographs*. Let (M, δ) be a pseudometric space and let us consider the class $\Im(M)$ of the fuzzy subsets of M with bounded support. Let $\underline{\delta}$ be a pseudometric on the interval [0, 1]. We denote by δ_{box} the *box-pseudometric* product in $M \times [0, 1]$ defined by setting

$$\delta_{box}\,[(x_1, y_1), (x_2, y_2)] = max\,\{\delta(x_1, x_2), \underline{\delta}(y_1, y_2)\},$$

for every $x_1, x_2 \in M$ and $y_1, y_2 \in [0, 1]$. For every $s \in \Im(M)$ the *hypograph* of s is the set

$$H(s) = \{(x, y) \in M \times [0, 1] / y < s(x)\}.$$

Since

$$s \subseteq s' \Leftrightarrow H(s) \subseteq H(s'),$$

for every $s, s' \in \Im(M)$, the map H: $\Im(M) \rightarrow P(M \times [0, 1])$ is an embedding from the lattice $\Im(M)$ into the lattice $P(M \times [0, 1])$. So, we can identify each fuzzy set s with its hypograph $H(s)$ and this suggests to define the distances between two fuzzy sets as the distances between the related hypographs:

$$d^H(s, s') = d(H(s), H(s')) =$$

$$= inf\{\delta_{box} [(x_1, y_1), (x_2, y_2)] / (x_1, y_1) \in H(s), (x_2, y_2) \in H(s')\}$$

and

$$D^H(s, s') = d(H(s), H(s')) =$$

$$= sup\{\delta_{box} [(x_1, y_1), (x_2, y_2)] / (x_1, y_1) \in H(s), (x_2, y_2) \in H(s')\}.$$

Proposition 3.5.3 *Let* Δ^H*:* $\Im(M)\times \Im(M) \rightarrow I(\mathbb{R}_0^+)$ *be a map defined by*

$$\Delta^H (s, s') = [d^H(s, s'), D^H(s, s')].$$

Then, $(F(M), \subseteq, \Delta^H)$ *is an* ***ISM****-space in which the weight of a fuzzy set s coincides with the diameter of its hypograph, i.e.* $p^H(s) =| \mathrm{H(s)} |$.

Let us stress that this space does not extend the previously defined space $(C^*, \subseteq, \Delta_p^*)$, as the following example shows:

Example 1. Let M the euclidean line and $s = \chi_X$ the characteristic function related to $X =\{(x, y)$: $0\leq x\leq 1, y=0\}$. Then $p^H(s) = \sqrt{2}$, while $p_1^*(\mathrm{s}) = 1$. Therefore, $\Delta^H(s, s)\neq \Delta_1^*(s, s)$ and so, $(\Im(M), \subseteq, \Delta^H)$ is not an extension of $(C^*, \subseteq, \Delta_1^*)$.

In order to construct the related metric space of points, recall that an abstraction process in $(\mathfrak{I}(M), \subseteq, \Delta^H)$ is a decreasing sequence $\langle s_n \rangle_{n\in\mathbb{N}}$ in $\mathfrak{I}(M)$ such that $lim_{n\to\infty}\, p^H(s_n) = 0$. Let us observe that in $(C^*, \subseteq, \Delta_1{}^*)$ there exist abstraction processes which fail to be abstraction processes in $(\mathfrak{I}(M), \subseteq, \Delta^H)$, i.e. $lim_{n\to\infty}\, p_1{}^*(s_n) = 0$ does not imply that $lim_{n\to\infty}\, p^H(s_n) = 0$, for some decreasing sequences $\langle s_n \rangle_{n\in\mathbb{N}}$, as in the following examples:

Example 2. Let M be the Euclidean line and $P_0 \in M$. The sequence of normal fuzzy sets $\langle \chi_{B_n(P)} \rangle_{n\in\mathbb{N}}$, where $\chi_{B_n(P)}$ is the characteristic function related to the closed ball $B_n(P_0)$ of radius $1/n$ and center P_0, is an abstraction process with respect to the space $(C^*, \subseteq, \Delta_1{}^*)$, but $lim_{n\to\infty}\, p^H(\chi_{B_n(P)}) = 1$.

Example 3. Let M be the interval $[0, 1]$. We can consider the sequence $\langle s_n \rangle_{n\in\mathbb{N}}$ in $\mathfrak{I}(M)$ defined by $s_n(P) = P^n$ for every $P \in M$. We obtain that $lim_{n\to\infty}\, p^H(s_n) = \sqrt{2}$, while $lim_{n\to\infty}\, p_1{}^*(s_n) = 0$.

Now we need to show that $AP(\mathfrak{I}(M), \subseteq, \Delta^H) \neq \varnothing$. So, let $P_0 \in M$ and, for every $n\in\mathbb{N}$, let us define the map

$$s_n^{P_0}(P) = \begin{cases} 1/n & if\ P \in B_n(P_0) \\ 0 & if\ P \notin B_n(P_0). \end{cases} \tag{3.10}$$

We call $s_n^{P_0}$ the *fuzzy set centered in* P_0. Clearly, the sequence $\langle s_n^{P_0} \rangle_{n\in\mathbb{N}}$ is an abstraction process with respect to the space $(\mathfrak{I}(M), \subseteq, \Delta^H)$. As usual, since $AP(\mathfrak{I}(M), \subseteq, \Delta^H) \neq \varnothing$, we can define a pseudometric distance on it:

$$\delta^H(\langle s_n \rangle_{n\in\mathbb{N}}, \langle s'_n \rangle_{n\in\mathbb{N}}) = lim_{n\to\infty} \Delta^H(s_n, s'_n),$$

and the induced metric distance:

$$\delta'^H([\langle s_n\rangle_{n\in\mathbb{N}}], [\langle s'_n\rangle_{n\in\mathbb{N}}]) = \delta^H(\langle s_n\rangle_{n\in\mathbb{N}}, \langle s'_n\rangle_{n\in\mathbb{N}}).$$

We denote by (M_H, δ'^H) the metric space associated to $(\mathfrak{I}(M), \subseteq, \Delta^H)$.

Proposition 3.5.4 *Let (M, δ) be a metric space and $\mathfrak{I}(M)$ be the class of the fuzzy subsets of M. Then (M, δ) is isometric to (M_H, δ'^H).*

Proof. Let us define the following map $\varphi : (M, \delta) \to (M_H, \delta'^H)$ such that

$$P \to [\langle s_n^P\rangle],$$

where s_n^P is the fuzzy set centered in P, defined as in (3.10). Let P, $Q \in M$; since, for $n\to\infty$, it results that $H(s_n^P) \to P$ and $H(s_n^Q) \to Q$, we have that

$$\delta'^H([\langle s_n^P\rangle], [\langle s_n^Q\rangle]) = lim_{n\to\infty}\Delta^H(s_n^P, s_n^Q) = lim_{n\to\infty}[d^H(s_n^P, s_n^Q), D^H(s_n^P, s_n^Q)]$$

$$= lim_{n\to\infty} [D(H(s_n^P), H(s_n^Q))] = \delta(P, Q).$$

Moreover, φ is injective, since it is an isometry defined on a metric space. Finally let $\langle s_n\rangle \in AP(\mathfrak{I}(M), \subseteq, \Delta^H)$, then $H(s_n)$ converges to a point $P \in M$, otherwise $lim_{n\to\infty}\, p_H(s_n)>0$. Therefore $[\langle s_n\rangle] = [\langle s_n^P\rangle]$, and so φ is surjective.

□

Let us observe that the resulting metric space of points is always isometric to the starting one, without requiring its completeness. Called $(\hat{M}, \hat{\delta})$ the completion of (M, δ), we can construct two spaces: $(\hat{M}_C^*, \hat{\delta}^*)$ isometric to $(\hat{M}, \hat{\delta})$ and (M_H, δ'^H) isometric to (M, δ).Therefore, if (M, δ) is not complete, we have two non isometric models $(\hat{M}_C^*, \hat{\delta}^*)$ and (M_H, δ'^H) But, if (M, δ) is complete then the two spaces are isometric.

3.6 Applications to rough sets through interval sets

In this section we define approximate distances between rough sets through approximate distances between interval sets. The theory of rough sets, developed since the early 1980s (see [36]), is an extension of the set theory, different from and complementary to other generalizations, such as fuzzy sets. It is useful to handle incomplete information. There are different approaches to interpret the theory of rough sets and all can be explained using the notions of lower and upper approximation in an approximation space (see [49]). So, let us introduce these notions.

Let U be a nonempty set and θ be an equivalence relation on U. The pair (U, θ) is called *approximation space* (see [14]). The objects in U are *approximated* in the sense that they can only be distinguished up to their equivalence classes while the objects in the same equivalence class are indistinguishable. The notion of *rough set* is based on that of approximation space (see [14] and [15]). It represents a situation in which we are not able to describe precisely a given subset of the universe, using the available information, that is, equivalence classes of the equivalence relation on the universe. Instead we can form a pair of approximations. More precisely, given an approximation space (U, θ) and $X \subseteq U$, denoted by $[x]_\theta$ the equivalence class of x modulo θ, for every $x \in U$, then

$$\overline{X} = \cup\{[x]_\theta : [x]_\theta \cap X \neq \varnothing\}$$

is the *upper approximation* of X, and

$$\underline{X} = \cup\{[x]_\theta : [x]_\theta \subseteq X\}$$

is the *lower approximation* of X. If $X \subseteq U$, is the extension of a property P, then

- $x \in \overline{X}$ means that x *possibly* has the property P,
- $x \in \underline{X}$ means that x *certainly* has the property P.

The *area of uncertainty* extends over $\overline{X} \setminus \underline{X}$.

The idea is that we are interested in X, but in many real-life situations the only information we have is that X is between the set $\underline{X}$ of all points that definitely have the property P and the set $\overline{X}$ of all points that may have the property P. So the available information about the unknown set X can be given by $(\underline{X}, \overline{X})$.

Let us observe that $\underline{X}$, and $\overline{X}$ represent respectively the closure and the interior of X in the topology τ_θ, induced on U by the relation θ, i.e. the topology having as a basis the equivalence classes modulo θ. In such a topology the class of closed sets coincide with the class of open sets.

Definition 3.6.1 Let (U, θ) be an approximation space. For every $X \subseteq U$, a rough set is a pair $(\underline{X}, \overline{X})$. X is called *θ-definable* if $\underline{X} = \overline{X}$.

It follows that if X is *θ-definable*, then it is either empty or a union of equivalence classes of θ and it is an open set in the topology τ_θ. In this case, the area of uncertainty is $\emptyset$.

In the collection $\mathfrak{R}$ of all rough sets in U we can define the order relation

$$(\underline{X}, \overline{X}) \leq (\underline{Y}, \overline{Y}) \text{ iff } \underline{X} \supseteq \underline{Y} \text{ and } \overline{X} \subseteq \overline{Y}.$$

The theory of rough sets can be related with the one of interval sets. Let us recall this notion.

Definition 3.6.2 Let A_1, A_2 be subsets of U. A *closed interval* set is the subset of $P(U)$

$$[A_1, A_2] = \{X \in P(U) \,/\, A_1 \subseteq X \subseteq A_2\}.$$

The set A_1 is called the lower bound and the set A_2 the upper bound. In the class of all closed interval sets $I(P(U))$, we consider the usual an inclusion relation. Obviously, we have:

$[X_1, X_2] \subseteq [Y_1, Y_2]$ iff $X_1 \supseteq Y_1$ and $X_2 \subseteq Y_2$.

Any rough set $(\underline{X}, \overline{X})$ produces an interval set $[\underline{X}, \overline{X}]$ in which the lower and upper bounds are θ-definable sets. In fact, there exists an homomorphism from rough sets algebra to interval sets algebra (see [49]). More precisely, if we consider the subalgebra of the interval sets with θ-definable lower and upper bounds, a rough set can be viewed as an interval set.

Let us assume that in U a pseudometric δ is defined and let us consider the related Hausdorff distance between sets. Then we can define two maps between interval sets by setting:

$$d_{I(P(U))}([X_1, X_2], [Y_1, Y_2]) = inf\{\delta_H(A, B) / X_1 \subseteq A \subseteq X_2 \text{ and } Y_1 \subseteq B \subseteq Y_2\}$$

and

$$D_{I(P(U))}([X_1, X_2], [Y_1, Y_2]) = sup\{\delta_H(A, B) / X_1 \subseteq A \subseteq X_2 \text{ and } Y_1 \subseteq B \subseteq Y_2\}.$$

Since it seems natural to assume that (U, δ) is bounded, $D_{I(P(U))}$ is always finite. In accordance we can define the interval valued map $\Delta_{I(P(U))}$: $I(P(U)) \times I(P(U)) \to I(\mathbb{R}_0^+)$ as:

$$\Delta_{I(P(U))}([X_1, X_2], [Y_1, Y_2]) = [d_{I(P(U))}([X_1, X_2], [Y_1, Y_2]), D_{I(P(U))}([X_1, X_2], [Y_1, Y_2])].$$

Proposition 3.6.1 *The space* $(I(P(U)), \subseteq, \Delta_{I(P(U))})$ *is an ISM-space, where the weight of an interval set* $p_{I(P(U))}([X_1, X_2])$ *is equal to* $\delta_H(X_1, X_2)$.

Proof. We observe only that

$$\begin{aligned} p_{I(P(U))}([X_1, X_2]) &= sup\{\delta_H(A, B) / X_1 \subseteq A \subseteq X_2 \text{ and } X_1 \subseteq B \subseteq X_2\} \\ &= sup\{e_\delta(A, B), e_\delta(B, A) / X_1 \subseteq A \subseteq X_2 \text{ and } X_1 \subseteq B \subseteq X_2\} \\ &= e_\delta(X_2, X_1) = \delta_H(X_2, X_1). \end{aligned}$$

□

According to what we stated regarding the relations between interval sets and rough sets, we can define in an analogous way an interval valued map between rough sets $\Delta_{\mathfrak{R}}$: $\mathfrak{R} \times \mathfrak{R} \to I(\mathbb{R}_0^+)$ as:

$$\Delta_{\mathfrak{R}}((\underline{X}, \overline{X}), (\underline{Y}, \overline{Y})) = [d_{\mathfrak{R}}((\underline{X}, \overline{X}), (\underline{Y}, \overline{Y})), D_{\mathfrak{R}}((\underline{X}, \overline{X}), (\underline{Y}, \overline{Y}))].$$

Thus, we have the following

Proposition 3.6.2 *The space* $(\mathfrak{R}, \leq, \Delta_{\mathfrak{R}})$ *is an ISM-space, in which the weight of a rough set is given by*

$$p_{\mathfrak{R}}(\underline{X}, \overline{X}) = \delta_H(\underline{X}, \overline{X}).$$

The *roughness* of $(\underline{X}, \overline{X})$ becomes as weaker as the lower and the upper approximations get closer. So, the weight $p_{\mathfrak{R}}$ indicates how much the approximation of X is rough, i.e. it provides a measure of the area of uncertainty, evaluating the distance between $\underline{X}$ and $\overline{X}$.

Evidently, if we consider the metric space associated to $(\mathfrak{R}, \leq, \Delta_{\mathfrak{R}})$, we denote by $(M_{\mathfrak{R}}, \delta^{\mathfrak{R}})$, the points in this space are just the exact sets, i.e. the open sets in the topology τ_θ, that represent complete information.

Thus we can define the map Φ: $(\tau_\theta, \delta_H) \to (M_{\mathfrak{R}}, \delta^{\mathfrak{R}})$ by

$$\Phi(X) = [\langle(\underline{X}_n, \overline{X}_n)\rangle_{n\in\mathbb{N}}],$$

where $X_n = X$ for each $n\in\mathbb{N}$.

Proposition 3.6.3 *The following hold*

i) Φ *is surjective and* $\delta^{\mathfrak{R}}(\Phi(X), \Phi(Y)) = \delta_H(X, Y)$, *for every* $X, Y \in \tau_\theta$;

ii) *if any open set in* τ_θ *is closed in the topology induced by the metric* δ, *then* Φ *is a bijection, i.e. the metric spaces* (τ_θ, δ_H) *and* $(M_{\mathfrak{R}}, \delta^{\mathfrak{R}})$ *are isometric.*

3.7 Approximate distances between E-rough sets

In this section we examine rough sets based on approximation spaces obtained considering a graded notion of equivalence.

Let $E:U\times U \to [0, 1]$ be a similarity with the $\wedge$ t-norm. E defines an order-reversing family of equivalence relations $(\theta_\lambda)_{\lambda\in[0, 1]}$, by its λ-cuts $C(E, \lambda) = \{(x, y)/ E(x, y)\geq\lambda\}$. Indeed we set

$$x\ \theta_\lambda\ y \Leftrightarrow (x, y) \in C(E, \lambda),$$

for every $x, y \in U$ and $\lambda \in [0, 1]$. For any $\lambda_1, \lambda_2 \in [0, 1]$, $\lambda_1 \leq \lambda_2$ entails that $\vartheta_{\lambda_1} \supseteq \vartheta_{\lambda_2}$, thus every class of ϑ_{λ_1} is a union of classes of ϑ_{λ_2}. The finest equivalence relation we can obtain is ϑ_1, while ϑ_0 is the universal relation $U\times U$.

Let us call *fuzzy approximation space* the pair (U, E). Given $X \subseteq U$, for each $\lambda \in [0, 1]$, the equivalence θ_λ defines a rough set $(\underline{X}_\lambda, \overline{X}_\lambda)$ we call *λ-level rough set.* We have an order-reversing family of rough sets $((\underline{X}_\lambda, \overline{X}_\lambda))_{\lambda\in[0,1]}$ and for $\lambda_1, \lambda_2 \in [0, 1]$, $\lambda_1 \leq \lambda_2$, it results that $\underline{X}_{\lambda_1} \subseteq \underline{X}_{\lambda_2} \subseteq \overline{X}_{\lambda_2} \subseteq \overline{X}_{\lambda_1}$.

For $\lambda \to 1$, i.e. for the equivalence relations tending to the finest one, in such a family of rough sets, the weights $p_{\mathfrak{R}}((\underline{X}_\lambda, \overline{X}_\lambda))$ decrease. If the similarity we consider is such that $E(x, y) = 1 \Leftrightarrow x = y$, the finest equivalence relation we can obtain is the identity relation and $\underline{X}_1 = \overline{X}_2$. In this case, for $\lambda \to 1$ the weights of the λ-level rough sets tend to zero and so the rough sets $(\underline{X}_\lambda, \overline{X}_\lambda)$ tend to the set X.

To evaluate the approximate distance $\Delta_{\mathfrak{R}}$ between two rough sets of the same family, let us observe that for $\lambda_1 \leq \lambda_2$ it results that $d_{\mathfrak{R}}((\underline{X}_{\lambda_1}, \overline{X}_{\lambda_1}), (\underline{X}_{\lambda_2}, \overline{X}_{\lambda_2}))= 0$, trivially. Since for any sets $A \subseteq B$, $\delta_H(A, B) = e_\delta(B, A)$, it results $D_{\mathfrak{R}}((\underline{X}_{\lambda_1}, \overline{X}_{\lambda_1}), (\underline{X}_{\lambda_2}, \overline{X}_{\lambda_2}))=e_\delta(\overline{X}_{\lambda_2}, \underline{X}_{\lambda_1}) \vee e_\delta(\overline{X}_{\lambda_1}, \underline{X}_{\lambda_2})$.

Definition 3.7.1 Given a similarity E: $U\times U \to [0, 1]$, we call *E-rough set* any family $((\underline{X}_\lambda, \overline{X}_\lambda))_{\lambda\in[0,1]}$ defined by a subset X of U.

We can extend the approximate distance between rough sets to an approximate distance between two E-rough sets. Indeed we set

$$d_{F_R}\,(((\underline{X}_\lambda,\ \overline{X}_\lambda))_{\lambda\in[0,1]},\ ((\underline{Y}_\lambda,\ \overline{Y}_\lambda))_{\lambda\in[0,1]}\,) = \int_0^1 d_{\Re}((\underline{X}_\lambda,\ \overline{X}_\lambda),\ (\underline{Y}_\lambda,\ \overline{Y}_\lambda))\ d\lambda$$

and

$$D_{F_R}\,(((\underline{X}_\lambda,\ \overline{X}_\lambda))_{\lambda\in[0,1]},\ ((\underline{Y}_\lambda,\ \overline{Y}_\lambda))_{\lambda\in[0,1]}\,) = \int_0^1 D_{\Re}((\underline{X}_\lambda,\ \overline{X}_\lambda),\ (\underline{Y}_\lambda,\ \overline{Y}_\lambda))\ d\lambda.$$

If we define an inclusion $\leq$ between E-rough sets as
$((\underline{X}_\lambda,\ \overline{X}_\lambda))_{\lambda\in[0,1]} \ \leq\ ((\underline{Y}_\lambda,\ \overline{Y}_\lambda))_{\lambda\in[0,1]} \ \Leftrightarrow\ (\underline{X}_\lambda,\ \overline{X}_\lambda) \ \leq\ (\underline{Y}_\lambda,\ \overline{Y}_\lambda)$ for every $\lambda\in[0, 1]$, we obtain

Proposition 3.7.1 *Let $F_{\Re}$ be the class of the E-rough sets and let Δ_{F_R} be defined by d_{F_R} and D_{F_R}. Then, the space $(F_{\Re},\ \leq,\ \Delta_{F_R})$ is an **ISM**-space.*

Proof. Properties of $d_{\Re}$ and $D_{\Re}$ are extended to d_{F_R} and D_{F_R} by properties of integrals.

□

3.8 Applications to clustering

Clustering is a process of distributing data (or objects) into groups called *clusters*, so that objects in the same cluster are more similar to each other than to objects in any other group. Sorting happens on the basis of similarities or, equivalently, distances between data. In order to group data, there are two ways of choosing distances. It is possible either to define a punctual distance between the initial data, at first, and then a distance between sets of grouped data, or to define a distance between sets of data and then to regard the punctual distance as a special case of this one. In this section, we choose to use a distance between sets, the approximate distance Δ defined in Proposition 3.4.1, extending a classical clustering procedure.

Let us see how, generally, a clustering technique works.

Let $X = \{x_1,\ldots, x_n\}$ be the set of n objects to be clustered into K clusters. The set $G = \{G_1,\ldots, G_K\}$ (called *clustering*) of the clusters must be a partition of X, i.e.

$$\bigcup_{i=1}^{K} G_i = X,\ \ G_i \cap G_j = \emptyset, \text{ for } i \neq j \ \text{ and } G_i \neq \emptyset.$$

Clustering algorithms are divided into two wide groups: hierarchical and optimization algorithms. Proceeding by either a series of successive mergers or a series of successive divisions, hierarchical algorithms produce a hierarchy of related clusterings which can be arranged in a tree-like structure known as *dendogram*. Instead, optimization techniques produce a single clustering which optimizes a pre-fixed criterion or objective function.

We focus on the *Agglomerative Hierarchical Technique*, whose characteristic is that a pair of clusters G_i, G_j is selected and merged at a time, whereby the number of clusters is reduced by one. Now we examine this method.

Agglomerative Hierarchical Clustering Procedure

Let us suppose to have a family of n data $X = \{x_1,\ldots, x_n\}$ to be clustered and let us assume that a distance d: $P(X)\times P(X) \to \mathbb{R}$ is defined.

Step 0

Set every data in its own cluster, i.e. set $K = n$ and $G_i = \{x_i\}$.

Step 1

Compute the matrix of the distances between the clusters $(d(G_i, G_j))_{i,j}$.

Step 2

Select a pair of clusters G_p and G_q such that

$$d(G_p, G_q) = min_{i\neq j}\, d(G_i, G_j).$$

Step 3

Merge the clusters selected in Step 2:

$$G_r = G_p \cup G_q$$

and set $K = K - 1$.

Step 4

If $K = 1$, stop, else go to Step 1 to update the matrix of the distances.

Every iteration of the algorithm yields as output a clustering; we obtain a hierarchy starting from the first clustering consisting of singletons up to the last clustering which is just the set X. Let us observe that Step 2 is not deterministic. So, some selection strategy is also necessary.

There are many different definitions we can choose for the distance and many different ways of updating it. For example, in the *single linkage method* the distance between two clusters is the minimum distance; the *complete linkage method* uses the maximum distance, whereas the *average linkage method* utilizes the distance between the centroids of two clusters. As another example, the distance used in [44] is the *Hausdorff* metric. In this case, since for the *excesses* it results

$$e_\delta(G_p \cup G_q, G_i) = e_\delta(G_p, G_i) \vee e_\delta(G_q, G_i)$$

$$e_\delta(G_i, G_p \cup G_q) = e_\delta(G_i, G_p) \wedge e_\delta(G_i, G_q,),$$

we can update the distances in the following way:

$$\delta_H(G_r, G_i) = \delta_H(G_p \cup G_q, G_i) = e_\delta(G_p \cup G_q, G_i)) \vee e_\delta(G_i, G_p \cup G_q)$$

$$= e_\delta(G_r, G_i) \vee e_\delta(G_i, G_r).$$

Since the Hausdorff distance is defined on sets and since it falls just into the interval whose end points are the infimum and supremum distances, as we showed in Proposition 3.4.2, it seems natural to utilize in the procedure above described the approximate distance Δ between subsets we defined in Proposition 3.4.1. In such a case, Δ provides the range in which the distance between two clusters can vary.

We start with the same initial situation in Step 0 and the same situation of computing the distances in Step 1.Namely, to update the distances, we have to calculate, for every $i \neq r$

$$\Delta(G_r, G_i) = \Delta(G_p \cup G_q, G_i) = [d(G_p \cup G_q, G_i), D(G_p \cup G_q, G_i)],$$

where

$$d(G_p \cup G_q, G_i) = d(G_p, G_i) \wedge d(G_q, G_i)$$

and

$$D(G_p \cup G_q, G_i) = D(G_p, G_i) \vee D(G_q, G_i)$$

by properties (3.7) e (3.8). Moreover, in Step 2 we have to select G_p and G_q in such a way that

$$\Delta(G_p, G_q) = min_{i \neq j} \Delta(G_i, G_j),$$

where the minimum is defined with respect to a suitable pre-order between intervals. As an example, we can define linear pre-orders on the set $I(\mathbb{R}_0^+)$, profiting by the usual order relation on $\mathbb{R}$. Indeed, let f be a map f: $I(\mathbb{R}_0^+) \rightarrow \mathbb{R}$, then we set

$$[u, v] \leq_f [u', v'] \text{ iff } f([u, v] \leq f([u', v']).$$

In such a way Step 2 becomes

Step 2

Select G_p and G_q in such a way that

$$f(\Delta(G_p, G_q)) = min_{i \neq j} f(\Delta(G_i, G_j)).$$

Obviously, we can choose the map f in different ways. By considering the first and second projections, we obtain the just exposed infimum and maximum distances based procedures. As an other example, we can use as f the map that associates each interval $[u, v]$ with its *middle point* $(u + v)/2$.

In the case in which the pre-order associated with f is not sufficient we can also introduce a second map g: $I(\mathbb{R}_0^+) \rightarrow \mathbb{R}$ in addition to f and set

$$[u, v] \leq_{fg} [u', v'] \text{ iff } \begin{cases} f([u,v]) < f([u',v']) \\ \quad\quad or \\ f([u,v]) = f([u',v']) \text{ and } g([u,v]) \leq g([u',v']). \end{cases}$$

By taking as f and g the first and the second projection, respectively, we obtain a total order, the so called l*exicographical order*. In such a case

$$[u, v] \leq_l [u', v'] \text{ iff } u < u' \text{ or } (u = u' \text{ and } v \leq v')$$

Generally, if a method of Agglomerative Clustering, with a distance δ satisfies the property

$$\delta(G_p \cup G_q, G_i) \geq \delta(G_p, G_i) \wedge \delta(G_q, G_i),$$

then, the method does not induce any reversal in the dendogram (see). Let us remark that the procedure utilizing Δ satisfies this property, with respect to the lexicographical order. In fact, it results

$$[d(G_p, G_i) \wedge d(G_q, G_i), D(G_p, G_i) \vee D(G_q, G_i)] \geq$$
$$[d(G_p, G_i) \wedge d(G_q, G_i), (D(G_p, G_i) \wedge D(G_q, G_i)) \vee D(G_p, G_i)]$$

i.e.

$$[d(G_p \cup G_q, G_i), D(G_p \cup G_q, G_i)] \geq [d(G_p, G_i), D(G_p, G_i)] \wedge [d(G_q, G_i), D(G_q, G_i)].$$

Let us observe that selecting two clusters G_p and G_q according to lexicographical order means that at first we check the minimum distances between the clusters and we individuate the class of the pairs which have the smallest minimum distance; successively, in such a class we select a pair which have the smallest maximum distance.

We conclude this section by observing that there are several further possibilities. As an example, we can define a (total) order relation between intervals checking first their middle points and then their widths, i.e.

$$[u, v] \leq_I [u', v'] \text{ iff } \begin{cases} (u+v)/2 < (u'+v')/2 \\ \qquad\qquad or \\ (u+v)/2 = (u'+v')/2 \;\; and \;\; w([u,v]) \leq w([u',v']). \end{cases}$$

Chapter 4

Fixed points, quasi-metrics and fuzzy orders

In this chapter, among the distances verifying weaker axioms than the ones of metrics, we focus on quasi-metics and on their dual notion of fuzzy order. In particular we take under consideration fixed points theory, both in a metric setting and in ordered sets, and its application to logic programming.

4.1 Preliminaries

Fixed point theory for operators in a lattice is a basic tool for formal logic. In fact, if we have a lattice L whose elements represent "pieces of information", in a logical apparatus it is usually possible to define an *immediate consequence operator* T. The order in L can be intended with respect to the informative content. In other words, saying that $x \leq y$ means that y represents more complete information than x. In accordance with this interpretation, given $x \in L$, $T(x)$ is the information we can obtain from x by one step of the inferential process, and the fixed points of T represent the deductively closed pieces of information. The least fixed point of T greater than or equal to x represents the theory generated by x (as an example, see [12] for crisp logic and [4] or fuzzy logic).

In particular, fixed point theory is very useful in logic programming. In such a field, the lattice L is the power set $P(B_P)$, where B_P is the Herbrand base associated with a program P, and T is the single-step operator, T: $P(B_P) \to P(B_P)$, associated with the program P. The fixed points of T are the Herbrand models for P and, therefore, the least fixed point of T is the least Herbrand model of P.

If the logic under consideration is monotone, in particular, if the program P is positive, T is a monotone operator, and therefore it is possible to apply the fixed point theorem of Knaster and Tarski, for ordered sets. Nevertheless, when the single-step operator T is not monotone, for instance if T is associated with a program containing negation, fixed point theorems for ordered set appear to be insufficient. In such case, it can be helpful to consider an approach metric in nature and to apply fixed point techniques in metric spaces, which are derived from Banach-Caccioppoli's contraction theorem (see [13], [37]). Another reason suggesting the opportunity to refer to metric spaces comes from fuzzy logic. Indeed, the process leading to a fuzzy set of consequences from a fuzzy set of hypotheses happens in a continuous environment. Such a process cannot finish by giving the exact output. Rather is an endless approximation of the ideal output. From here the need arises to define someway the notion of "approximation". This is possible only in a metric setting.

On the other hand, fixed point theory in ordered sets and fixed point theory in metric spaces can be unified. Indeed, the notion of fuzzy order allows us to extend simultaneously both the metric notions and the ones of ordered sets theory. The notion of fuzzy order is dual to the one of quasi-metric, which also combines the basic usefulness of metrics in measuring the distance between two objects and the advantages of order in order-theoretic arguments, (see [37], [39], [43]). In accordance with this duality, it is possible to demonstrate a theorem simultaneously generalizing the fixed point theorem of Knaster and Tarski for ordered structures and the theorem of Banach- Caccioppoli for metric spaces.

4.2 Operators

We start with some hints of the theory of operators in the lattice of all the subsets of a given nonempty set. Afterwards we extend the related notions to any complete lattice.

Let S be a nonempty set and let us denote by $P(S)$ the lattice of all the subsets of S. We call *operator* in S any map J from $P(S)$ into $P(S)$.

Definition 4.2.1 An operator $J{:}P(S)\rightarrow P(S)$ is called *closure operator* if, for every X, $Y \in P(S)$, it satisfies:

i) $X \subseteq Y \Rightarrow J(X) \subseteq J(Y)$;

ii) $X \subseteq J(X)$;

iii) $J(J(X)) = J(X)$.

An *almost closure operator*, briefly *a-c-operator*, is an operator J satisfying *i)* and *ii)* of the previous definition.

Let us enunciate some important properties for operators. We can introduce the notion of *compactness*, which is a very useful property, since it corresponds to the notion of a finite "construction process".

Definition 4.2.2 Let J: $P(S) \rightarrow P(S)$ be an operator in S. Given $X\in P(S)$, J is *compact* if for every $x \in J(X)$, there exists a finite subset X_f of X such that $x \in J(X_f)$, that is

$$J(X)=\cup\{J(X_f) \,/\, X_f \textit{ finite subset of } X\}.$$

Now let us refer, for the above introduced notions, to any lattice. Let us denote, in the following, by L a complete lattice whose minimum and maximum we denote by 0 and 1, respectively.

Definition 4.2.3 An operator J: $L \to L$ is called a *closure operator* if, for every x, $y \in L$

i) $x \le y \Rightarrow J(x) \le J(y)$ (monotony)

ii) $x \le J(x)$ (inclusion)

iii) $J(J(x)) = J(x)$ (idempotence).

Since the notion of finite subset is not defined in a generic lattice L, we have to search for a different notion of compactness. In order to do this, let us introduce the notion of directed class. A nonempty class $C = (x_i)_{i \in I}$ of elements of L is called *directed* if

$$\forall\, i, j \in I,\ \exists\, h \in I \text{ such that } x_i \le x_h \text{ and } x_j \le x_h.$$

The chains are typical examples of directed classes. We say that $z = Sup(C)$ is *the limit of* C and we write $z = limC$. If J is an order-preserving operator and C is a directed family, then also its image $J(C) = \{J(x) / x \in C\}$ is directed, obviously.

Definition 4.2.4 An operator J is called *continuous* if it is order preserving and, for every directed class C,

$$J(limC) = limJ(C).$$

A continuous closure operator is also called an *algebraic closure operator*.

It is possible to prove that if L is the lattice $P(S)$ of all the subsets of a given set S, then J is continuous if and only if J is compact.

4.3 Fixed points in ordered sets

Let H be a continuous a-c-operator. The following famous theorem states that for every $x \in L$, the least fixed point of H greater or equal to x, $c(H)$, is given by $Sup_{n \in \mathbb{N}} H^n$.

Theorem 4.2.1 (*Fixed-point Theorem*) *Let H be a continuous a-c-operator. Then*

$$c(H) = Sup_{n\in\mathbb{N}}H^n.$$

Proof. We have to prove that, for every $x\in L$, $Sup_{n\in\mathbb{N}}H^n(x)$ is the least fixed point of H greater than or equal to x. Now, since $H(x) \geq x$ we have also that $H^{n+1}(x) \geq H^n(x)$ for every n, and hence the family $(H^n(x))_{n\in\mathbb{N}}$ is directed. Since H is continuous

$$H(Sup_{n\in\mathbb{N}}H^n(x)) = Sup_{n\in\mathbb{N}}H^{n+1}(x) = Sup_{n\in\mathbb{N}}H^n(x)$$

and $Sup_{n\in\mathbb{N}}H^n(x)$ is a fixed point for H greater than or equal to x. If y is any fixed point such that $y \geq x$, then for every $n\in\mathbb{N}$, $y = H^n(y) \geq H^n(x)$ and hence $y \geq Sup_{n\in\mathbb{N}}H^n(x)$. This proves that $Sup_{n\in\mathbb{N}}H^n(x)$ is the least fixed point of H greater or equal to x.

□

Let us consider now some applications of the theory of closure operators to the logic. We define an *abstract deduction system* as a pair (L, D) where L is a complete lattice and D a closure operator in L. The elements in L are called *pieces of information* and D the *deduction operator*. A *theory* is defined as a fixed point of D, i.e., as a piece of information τ such that $\tau \geq D(\tau)$. So the theories are the deductively closed pieces of information.

For example, the classical first order logic is an abstract logic in which the pieces of information are the sets of formulas (systems of axioms), $D(x)$ is the set of formulas we can derive from x, a theory τ is a set of formulas containing the logical axioms and closed under the inference rules .

A deduction operator D is usually obtained starting from a suitable set A of logical axioms and a suitable set of inference rules. Namely, let us denote by $J(X)$

the set of the formulas that can be obtained by one application of the inference rules to the formulas in X, and let us set

$$H(X) = J(X) \cup A \cup X,$$

i.e., $\alpha \in H(X)$ if: either α is obtained by applying an inference rule to the formulas in X, or α is a logical axiom, or α is an element in X.

It is immediate that H is an almost closure operator and $D(X)=\cup H^n(X)$. For a natural number n, $H^n(X)$ represents the set of the formulas we can obtain from X by an n-step inferential process.

Now, we consider the application of closure operators to a particular kind of theories, the programs. Given a first order language $\mathcal{L}$ and a program P, it is possible to associate to P a closure operator, the *single step operator*, which allows to obtain the Herbrand models of the program. These notions can be applied in the logic programming field.

4.4 Logic programming by fixed points approach

Logic programming concerns both logic and programming, so several kind of semantics have been developed for it. In this context we are interested to the application of fixed points theory to logic programming.

Let P be a definite program and B_P the Herbrand basis associated with it. Then, it is possible to introduce a continuous map on the complete lattice $P(B_P)$ and it is possible, by means of this map, to have a fixed point characterization of the least Herbrand model of P.

In order to examine the case of logic programming with negation, we prefer to consider $\{false, true\}^{B_P}$, instead of $P(B_P)$. In accordance we give the following

Definition 4.4.1 A (classical) *valuation* is a mapping v: $B_P \rightarrow \{false, true\}$ from the set of ground atoms to the set of classical truth values.

Indeed, it is common in the logic programming literature to identify a valuation v with the set of ground atoms which are true by v. The standard approach in logic programming is to take *false* as the default. It is possible to assign to the space $\{false, true\}$ the *truth-based* ordering $\leq_t$, which asserts $false \leq_t true$. This ordering is pointwise extended to valuation: $v_1 \leq_t v_2$ if and only if $v_1(A) \leq_t v_2(A)$, for every ground atom A. Moreover, it is possible to define a *single-step* operator for a program P, in the usual way (see [16]). It is a map T_P, from valuations to valuations, such that $T_P(v)$ makes a ground atom A true in the case that A is the head of a ground instance of some clause in P, and v makes the body of that ground instance *true*. Formally, given a program P, we have first to define P^* as the set associated to P, constructed by: taking all the ground instances of members of P, replacing the possible clauses $A\leftarrow$ with empty body with $A\leftarrow$ *true*, adding $A\leftarrow$*false*, if the ground atom A is not the head of any member of P^*. Then we can give the following

Definition 4.4.2 A single-step operator for a program P is a map T_P: $\{false, true\}^{B_P} \rightarrow \{false, true\}^{B_P}$, defined by $T_P(v)=w$, where, for a ground atom A, w is the unique valuation determined by the following:

(i) $w(A) = true$ if there is a ground clause $A \leftarrow B_1, \ldots, B_n$ in P^* with head A such that $v(B_1) = true$, and …, and $v(B_n) = true$.

(ii) $w(A) = false$ otherwise.

What we would like to obtain is a fixed point for the single-step operator, that is, a valuation that the program cannot edit. For any definite program, the associated single-step operator is monotone, i.e.

$$v_1 \leq_t v_2 \text{ implies } T_P(v_1) \leq_t T_P(v_2).$$

Since it is possible to verify that T_P is continuous, in this case, the well-known Knaster-Tarsky Theorem assures that the single step operator has smallest (and largest) fixed points. The smallest fixed point of T_P is the smallest Herbrand model.

4.5 Non monotone logics and fixed points

Let us now examine one of the most, at the same time controversial and worthwhile, addition to the basic logic programming mechanism: the negation. In a simple database language, negation is not a problem: it is possible to report the facts that either an item is in a database, or it is not. But if a system is built on classical first-order logic, negation can be a serious problem, (see [16]). A weaker version of negation can be used: *negation as failure*, that is, one concludes *not X* if *X* is not a logical consequence of a program. Negation as failure is essentially non monotonic: if *X* is not a consequence of the program, we say that *not X* is a conclusion, but if *X* is added to the program, then the conclusion *not X* must be removed. The addition of new axioms can decrease the set of theorems that previously held.

For a general logic program *P* the definition of the single-step operator could be extended from the one of definite programs, essentially requiring that $v(\textit{not } X)$ has the truth value $\neg v(X)$. Nevertheless, this is not a satisfactory way to proceed; indeed, since the presence of negations destroys monotonicity, the existence of smallest and biggest fixed points is no longer guaranteed. Moreover, it can happen that also for a very simple program, such as $P \leftarrow \textit{not } P$, it does not exist any fixed point. So, it seems necessary to bring some changes to the above described approach.

A possibility examined in literature is to consider *partial valuation*, or *three-valued* valuations.

Definition 4.5.1 A *partial valuation* is a mapping v: $B_P \rightarrow \{\perp, false, true\}$ from the set of ground atoms to the set $\{\perp, false, true\}$.

Once again it is common in the literature to work with sets, rather than with maps. The meaning is that if it is not possible to deduce if the value of a ground atom A is true or false, we take the value for A *undefined*, or $\perp$. In this case, the approach is to take $\perp$ as the default. More precisely, considering the smallest and the biggest fixed points of the single-step operator associated to a logic program, if they agree on one of the classical truth value for a ground atom A, then this value is taken to be the value of A. Otherwise, the value of A is undefined.

In the space $\{\perp, false, true\}$ it is possible to introduce the *knowledge-based* ordering $\leq_k$, that is an ordering based on the "degree of information" instead of the "degree of truth". It asserts $\perp \leq_k false$ and $\perp \leq_k true$. Again this ordering is pointwise extended to partial valuations, i.e. $v_1 \leq_k v_2$ if and only if $v_1(A) \leq_k v_2(A)$, for every ground atom A.

Now it is possible to define a new single-step operator Φ_P, from partial valuations to partial valuations, associated with a program P. In this case we need to specify when both the values *true* and *false* are assigned, otherwise $\perp$ is the default. So we have:

Definition 4.5.2 Given a program P, the single-step operator Φ_P is a map Φ_P: $\{\perp, false, true\}^{B_P} \rightarrow \{\perp, false, true\}^{B_P}$, defined by $\Phi_P(v)=w$, where, for a ground atom A, w is the unique partial valuation determined by the following:

(i) $w(A) = true$ if there is a general ground clause $A \leftarrow B_1, \ldots, B_n$ in P^* with head A such that $v(B_1) = true$, and …, and $v(B_n) = true$.

(ii) $w(A) = false$ if for every general ground clause $A \leftarrow B_1, \ldots, B_n$ in P^* with head A, $v(B_1) = false$, or …, or $v(B_n) = false$.

(iii) $w(A) = \perp$, otherwise.

In such a case, for a general program P, the operator Φ_P is monotone with respect to the ordering $\leq_k$. Since $\{\perp, \mathit{false}, \mathit{true}\}$ does not result a complete lattice, the Knaster-Tarski theorem cannot help us, but the algebraic structure is rich enough to ensure the existence of smallest fixed point (though not biggest) for monotone maps. Sometimes the three-valued approach results natural, but there are cases in which it is quite awkward. So in some situations it is more advantageous to introduce different techniques based on metric spaces, which result, when applicable, simpler, also from a computational point of view.

4.6 Metric methods for logic programming

The existence of a model for a logic program is generally established by lattice-theoretic arguments, as we saw. But it is often possible to use metric methods instead. In this section we examine some of these methods useful to handle general logic programs and some problems arising from them. We refer again to the approach based on fixed points theory. So, the aim is to find a fixed point for the single-step operator associated with a program.

One of the most significant metric approach is that of Fitting (see [16], [17]), who proved the existence of fixed points of the immediate consequence operator for some non-positive programs applying the Banach Contraction Theorem, as a replacement for the Knaster-Tarski Theorem. In this application it turned out that the metric defined was actually an ultrametric. More precisely he involved the notion of level mappings.

Definition 4.6.1 A *level mapping* for a program P is a function l: $B_P \to \mathbb{N}$ from ground atoms to natural numbers. We say that $l(A)$ is the level of the ground atom A.

Given $n \in \mathbb{N}$, we denote by l_n the set $l^{-1}(n)$ of all the atoms of level n.

It is possible to define a metric, between (classical) valuations, associated with a level mapping.

Definition 4.6.2 Let v and w be valuations and let l be a level mapping. Let us define the associated metric d as $d(v, w) = 0$, if $v = w$; otherwise, $d(v, w) = 1/2^n$, where v and w differ on some ground atom of level n, but agree on all ground atoms of lower levels.

It is quite a routine to show that d results an ultrametric. Moreover, it results that the space of valuations, using such a metric, based on a level mapping, is a complete metric space.

It is possible to have many possible metrics on the set of valuations, since every level mapping determines one. The aim is to find a metric with respect to which the single-step operator T_P of a program is a contraction: so it will be possible to resort to the well-known Banach Contraction Theorem to find fixed points. Indeed, by this Theorem, if T_P is a contraction on the space of valuations, it has a unique fixed point. Let us stress that this reasoning often holds also if we define a metric on the three-valued valuations space, in other words, also the Φ_P operator often results a contraction.

Sometimes, as we have illustrated, the three-valued single step operator Φ_P results monotonic with respect to some orderings and, then, by the usual lattice-theoretic approach, it is possible to prove that it converges to a fixed point. Nevertheless, this operator may need in general more than ω steps to rich its fixed point, whereas the Banach theorem, when it is applicable, gives convergence in ω steps.

Now let us report some example (see[16], [17]) of simple logic programs in order to show how the described metric approach works.

Example 1 – *P1*

$even(0) \leftarrow$

$even(s(x)) \leftarrow \neg\, even(x)$

We are considering numbers represented as numerals, a constant symbol 0, a successor function symbol *s*.

Let us represent some of the first steps of this program in explicit form:

$T_{P1}(\varnothing)=B_P$

$T^2{}_{P1}(\varnothing)=\{even(0)\}$

$T^3{}_{P1}(\varnothing)=\{even(0), even(s^2(0)), even(s^3(0)), even(s^4(0)), even(s^5(0)), \dots \}$

$=B_P \setminus \{even(s(0))\}$

$T^4{}_{P1}(\varnothing)=\{even(0), even(s^2(0))\}$

$T^5{}_{P1}(\varnothing)=\{even(0), even(s^2(0)), even(s^4(0)), even(s^5(0)), \dots \}$

$=B_P \setminus \{even(s(0)), even(s^3(0))\}$

…

It is possible to define a level mapping simply by setting $l(even(s^n(0))) = n$. Then, we can define the rising ultrametric by: $d(v, v) = 0$ and $d(v, w) = 1/2^n$, where n is such that $v(even(s^n(0))) \neq w(even(s^n(0)))$, but $v(even(s^k(0)))= w(even(s^k(0)))$, for every $k<n$. With such a distance the space of valuations results a complete ultrametric space and T_{P1} is a contraction. In fact, we can easily observe that if the valuations v and w agree on $even(s^k(0))$, then $T_{P1}(v)$ and $T_{P1}(w)$ will agree on $even(s^{k+1}(0))$. So, if the distance between v and w is $1/2^n$, then the distance between $T_{P1}(v)$ and $T_{P1}(w)$ will be $1/2^{n+1}$. Thus, $d(T_{P1}(v), T_{P1}(w)) \leq \frac{1}{2} d(v, w)$, and, by the Banach Theorem, T_{P1} has a unique fixed point.

Example 2 – *P2*

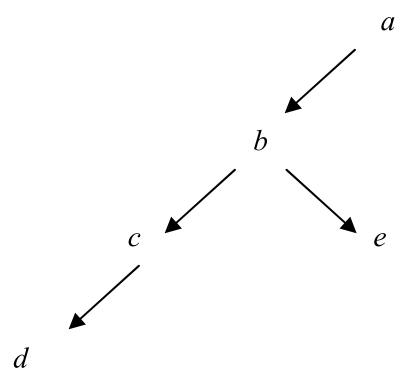
Let us suppose to have a game, with positions denoted by constants *a*, *b*, … and let us assume that impossibility of moving for a player means loosing. The program

starts with the list of all the legal moves of the game $move(a, b)$, $move(c, d)$, ….
We can write the program as follows:

$move(a_i, a_j) \leftarrow$ (for all legal moves a_i to a_j)

$win(x) \leftarrow move(x, y), \neg win(y)$

To show explicitly some steps of this programs, we examine a simple case. Let us suppose to have $move(a, b)$, $move(b, c)$, $move(b, e)$, $move(c, d)$:



In such a case it would be:

$T_{P2}(\varnothing)=\{move(a, b), move(b, c), move(b, e), move(c, d)\}$

$T^2_{P2}(\varnothing)=\{move(a,b), move(b,c), move(b, e), move(c, d), win(a), win(b), win(c)\}$

$T^3_{P2}(\varnothing)=\{move(a, b), move(b, c), move(b, e), move(c, d), win(b), win(c)\}$

$T^4_{P2}(\varnothing)=\{move(a, b), move(b, c), move(b, e), move(c, d), win(b), win(c)\}$.

In general we can say that if we assume that the program has no loops, the associated single-step operator has a unique fixed point. Indeed, it is possible to define a level mapping by setting $l(move(a, b))=1$, for every a, b, and $l(win(p))$ equal to the height of the tree having p at the root and each node labelled with a position, in such a way that the children of a node are labelled with the positions reachable in one move from the position labelling the parent. The finiteness of the game trees is guaranteed by the assumption that the program has no loops. It is easy to verify that T_{P2} results a contraction, with respect to the ultrametric corresponding to this level mapping.

Despite the success of metric methods with simple programs, various problems remain. Indeed, the applications of metric techniques has been extended in several directions. For example, Seda (see [42], [43]) examined a process to find fixed points of immediate consequence operator, utilizing quasi-metrics. This approach is carried out in order to overcome some restrictions of the previous approaches and it results much more general. Indeed, by quasi-metrics, characterized, as we saw in Chapter 1, by the lack of symmetric property, it is possible to combine both the lattice-theoretic approach and the metric one and, in a sense, unify fixed points theory in ordered sets and fixed points theory in metric spaces.

Also in this approach the metric functions are defined by means of level mappings l, with the property for l_n to be a finite set for each n. Besides, given a level mapping l, Seda defines the function *rank*, r: $B_D \rightarrow \mathbb{N}$, where B_D is the set of all finite subsets of B_P, by setting $r(I)=max_{A\in I}(l(I))$, for nonempty $I\in B_D$, and $r(\emptyset)=0$. Equivalently, we can refer to the valuations, as usual. So we denote by B_D the set of all valuations v such that $v(A)$ is true only for a finite set of atoms. In such a case, the definition of $r(v)$ is obvious. We have

Definition 4.6.3 Let u_1 and u_2 be valuations and let l be a level mapping and r be the corresponding rank function. Let us define the associated metric d as $d(u_1, u_2)=inf\{1/2^n \,/\, v \le u_1 \Rightarrow v \le u_2$ holds for every valuation v with $r(v)\le n\}$.

The so defined distance d results to be a quasi-ultrametric and the space of valuations with d results complete and totally bounded. If the single-step operator results non-expansive, it is possible to refer to the Rutten-Smyth theorem for quasi-metrics. But, differently from the previous case, now the single-step operator associated with a program P does not always result non-expansive. As an example let us take up the Program P1 of Example 1:

$even(0) \leftarrow$

$even(s(x)) \leftarrow \neg\, even(x)$

We define a level mapping in the same way, $l(even(s^n(0))) = n$, but we have now that T_{P1} is not non-expansive with respect to the corresponding quasi-metric. Indeed, let us consider u_1 and u_2 corresponding to $I_1=\{even(0), even(s(0))\}$ and $I_2=\{even(0), even(s(0)), even(s^2(0))\}$, respectively. If we apply the single-step operator, we obtain that $T_{P1}(u_1)$ corresponds to $\{even(0), even(s^3(0)), even(s^4(0)), \ldots\}$ and $T_{P1}(u_2)$ corresponds to $\{even(0), even(s^4(0)), even(s^5(0)), \ldots\}$. Therefore, we have $d(u_1, u_2)=0$ (since $I_1 \subseteq I_2$), but $d(T_{P1}\,(u_1), T_{P1}(u_2))=2^{-2}$. In such a case, we can solve the problem by considering $T^n{}_{P1}(\emptyset)$: it is possible to verify that $d(T^n{}_{P3}(\emptyset), T^{n+1}{}_{P3}(\emptyset)=0$, if n is even, and $d(T^n{}_{P3}(\emptyset), T^{n+1}{}_{P3}(\emptyset)=2^{-n+1}$, if n is odd. Thus, $T^n{}_{P3}(\emptyset)$ is a Cauchy sequence and it converges to a fixed point.

It is possible to find programs whose single-step operators are never non-expansive, for any choice of the level mapping and corresponding distance (see, for instance [43]).

4.7 Fixed point theorems for fuzzy orders

We have seen how the notion of quasi-metric is involved in some problem regarding the search of fixed points. Due to what suggested by Valverde in [45] and examined in Chapter Chapter 1, the notion of fuzzy order is a dual one of the notion of quasi-metric distance. Then it should be interesting to investigate the possibility of a fixed points theory in the rather general framework of the theory of fuzzy orders.

Definition 4.7.1 Let M be a set. Given a map $f : M \to M$ and a fuzzy relation R we say that $x \in M$ is a *fixed point* for f (with respect to R) provided that

$$R(x, f(x)) = R(f(x), x) = 1.$$

In the case that R is a $*-$fuzzy order, x is a fixed point if and only if $f(x) = x$ (see Definition 1.4.1).

Now we give the dual notion of *forward Cauchy* sequence in a quasi-metric space (see, for example, [43]), by binary relations and in particular by fuzzy orders.

Proposition 4.7.1 *A sequence* $(x_n)_{n\in\mathbb{N}}$ *in a set M is said to be forward Cauchy if, for every* $0 \le \varepsilon < 1$*, there exists a natural number* $\underline{n}$ *such that* $R(x_n, x_m) \ge \varepsilon$ *whenever* $m \ge n \ge \underline{n}$.

Let us remark that when R is a fuzzy order, a sequence is forward Cauchy if and only if for each $0 < \varepsilon < 1$, there exists a natural number $\underline{n}$ such that $R(x_{n+1}, x_n) \ge \varepsilon$ for all $n \ge \underline{n}$.

Definition 4.7.2 Let $(x_n)_{n\in\mathbb{N}}$ be a forward Cauchy sequence. We say that $\ell \in M$ is a *limit* of $(x_n)_{n\in\mathbb{N}}$, if, for every $x \in M$, we have

$$R(\ell, x) = lim R(x_n, x).$$

We say that the forward Cauchy sequence $(x_n)_{n\in\mathbb{N}}$ *converges to* ℓ in M and we write $lim x_n = \ell$.

The structure (M, R) is called *complete* if every forward Cauchy sequence converges to a limit.

The proposed convergence depends on the convergence defined in $\mathbb{R}$, so it inherits lots of properties of the convergence in $\mathbb{R}$. As an example, if $(x_n)_{n\in\mathbb{N}}$ is a forward Cauchy sequence and ℓ is a limit of the sequence, an extract sequence of $(x_n)_{n\in\mathbb{N}}$ converges to the same limit ℓ.

For instance, let us assume that R is a partial order $\leq$. A sequence $(x_n)_{n\in\mathbb{N}}$ is forward Cauchy if and only if $\exists\, N,\ \forall n \geq N,\quad x_n \leq x_{n+1}$, i.e. if and only if it is "eventually a chain". Moreover, the statement $\lim x_n = \ell$ is equivalent to

$$\forall x \in M \quad (\ell \leq x \iff \exists m\ \forall n \geq m,\ \ x_n \leq x).$$

In particular, if x_n is order-preserving, then $lim x_n = \ell$ if and only if $\ell = Sup\{x_n \,/\, n \in \mathbb{N}\}$.

Proposition 4.7.2 *Let R be a $*$-fuzzy preorder. Then, two limits of a given sequence $(x_n)_{n\in\mathbb{N}}$ are similar. If R is a $*$-fuzzy order, then limit is unique.*

Proof. Let us assume that $lim x_n = \ell$ and $lim x_n = \ell'$. Then, by definition, $R(\ell, x) = lim R(x_n, x)$ and $R(\ell', x) = lim R(x_n, x)$ for every $x \in M$. In particular, by setting $x = \ell$, $1 = R(\ell, \ell) = lim R(x_n, \ell) = R(\ell', \ell)$ and, by setting $x = \ell'$, $1 = R(\ell', \ell') = \lim R(x_n, \ell') = R(\ell, \ell')$. Then $1 = R(\ell', \ell) = R(\ell, \ell')$ and ℓ is similar with ℓ'. Trivially, when R is a $*-$fuzzy order, by the antisymmetry limits coincide.

□

Let us give, now, some other definitions which are dual of some well-known notions in metric spaces theory. Let M be a set and f: $M \to M$ be a map.

Definition 4.7.3 Let R be a $*-$fuzzy preorder. The map f is called *continuous* if from $lim x_n = \ell$ it follows $lim f(x_n) = f(\ell)$, for every forward Cauchy sequence $(x_n)_{n\in\mathbb{N}}$ in M.

Obviously, when M is a partially ordered set, f is continuous if and only if it preserves upper bounds of chains.

Definition 4.7.4 We say that f is *non-expansive* if for every $x, y \in M$,

$$R(f(x), f(y)) \geq R(x, y).$$

Let us observe that if R is a $*-$similarity, then a non-expansive map is a function "compatible" with the $*-$similarity R. If R is a $*-$fuzzy preorder, then a non-expansive map is in a sense an order-preserving map.

Definition 4.7.5 We say that f is *contractive* if there exists $c > 1$ such that

$$(R(f(x), f(y)))^{c} \geq R(x, y).$$

In other terms, a contraction is a map that increases the similarity-degree between elements.

The following theorem is a dual version of the first part of Rutten-Smyth theorem regarding quasi-metrics, so it can be viewed as a unification of both metric notions and the ones of ordered sets theory.

Theorem 4.7.1. *Let R be a $*-$fuzzy preorder such that (M, R) is complete and let f: $M \to M$ be a non-expansive continuous map such that $R(x, f(x)) = 1$, for a suitable $x \in M$. Then f has a fixed point.*

Proof. Let us consider the sequence $(x, f(x), f^2(x),....)$. Since f is non-expansive, such a sequence is a forward Cauchy one. Indeed, trivially we have that

$$1 = R(x, f(x)) \leq R(f(x), f^2(x)) \leq ... \leq R(f^n(x), f^{n+1}(x)),$$

and so, by $*$-transitivity, it results, for every $m \geq n$,

$$R(f^n(x), f^m(x)) \geq R(f^n(x), f^{n+1}(x)) * ... * R(f^{m-1}(x), f^m(x)) \geq$$

$$\geq 1 * ... * 1 = 1$$

Moreover, from the hypothesis of completeness of (M, R), it follows that the sequence $(f^n(x))_{n\in\mathbb{N}}$ has a limit ℓ . Also, since f is continuous, we have that $f(\ell)$ is a limit of $(f^{n+1}(x))_{n\in\mathbb{N}}$, and therefore of $(f^n(x))_{n\in\mathbb{N}}$. Finally, since R is a $*$−fuzzy preorder, limits are similar, i.e. $R(f(\ell), \ell) = 1 = R(\ell, f(\ell))$. Then ℓ is a fixed point for f.

□

The following theorem is dual of fixed points theorems with contractive maps in metric settings and the proof makes use of some transpositions from the Banach contraction principle.

Theorem 4.7.2. *Let $*$ be a t-norm greater than or equal to the usual product. Let R be a $*$−fuzzy order such that (M, R) is complete and $f: M \to M$ be a continuous and contractive map. Then f has a unique fixed point.*

Proof. Let us observe that it is sufficient to prove the theorem with the *t*-norm of the product $\cdot$. Indeed, if $*$ is a *t*-norm greater or equal to the product, then the $*$−transitivity implies the $\cdot$−transitivity: $R(x, z) \geq R(x, y)*R(y, z) \geq R(x, y) \cdot R(y, z)$. Therefore, any $*$−fuzzy order is a $\cdot$−fuzzy order. Let x_0 be any element of M, and let $(x_n)_{n\in\mathbb{N}}$ be the sequence defined as follows: $x_1=f(x_0)$, $x_2=f(x_1)$, …, $x_{n+1}=f(x_n)$, ...

Let us prove that this sequence is forward Cauchy. Let us observe that by hypothesis there exists c, $0 < c < 1$, such that:

$$R(x_1, x_2) = R(f(x_0), f(x_1)) \geq (R(x_0, x_1))^{c} = (R(x_0, f(x_0)))^{c}$$

$$R(x_2, x_3) = R(f(x_1), f(x_2)) \geq (R(x_1, x_2))^{c} \geq (R(x_0, x_1))^{c^2} = (R(x_0, f(x_0)))^{c^2}$$

$$\ldots\ldots.$$

$$R(x_n, x_{n+1}) = R(f(x_{n-1}), f(x_n)) \geq (R(x_{n-1}, x_n))^{c} \geq (R(x_0, x_1))^{c^n} = (R(x_0, f(x_0)))^{c^n}\ .$$

Then, we have that

$$R(x_n, x_{n+r}) \geq R(x_n, x_{n+1}) \cdot R(x_{n+1}, x_{n+2}) \cdot \ldots \cdot R(x_{n+r-1}, x_{n+r}) \geq$$

$$\geq (R(x_0, f(x_0)))^{c^n} \cdot (R(x_0, f(x_0)))^{c^{n+1}} \cdot \ldots \cdot (R(x_0, f(x_0)))^{c^{n+r-1}} =$$

$$= (R(x_0, f(x_0)))^{c^n + c^{n+1} + \ldots + c^{n+r-1}}.$$

Let us observe that

$$c^n + c^{n+1} + \ldots. + c^{n+r-1} = \frac{c^n - c^{n+r-1}}{1-c} = \frac{c^n\left(1-c^{r-1}\right)}{1-c} \leq \frac{c^n}{1-c}.$$

Let us set $d = R(x_0, f(x_0))$. For every ε, such that $0 < \varepsilon < 1$, it is $log_d(\varepsilon) > 0$ and since $lim_{n\to\infty} \dfrac{c^n}{1-c} = 0$, there exists a natural number n_0 such that $\dfrac{c^n}{1-c} \leq log_d(\varepsilon)$ for any $n \geq n_0$. Therefore, we have $c^n + c^{n+1} + \ldots. + c^{n+r-1} \leq log_d(\varepsilon)$, for any $n \geq n_0$.

Consequently, for any $n \geq n_0$ and $r \in \mathbb{N}$,

$$R(x_n, x_{n+r}) \geq R(x_n, x_{n+1}) \cdot R(x_{n+1}, x_{n+2}) \cdot \ldots \cdot R(x_{n+r-1}, x_{n+r}) \geq$$

$$\geq d^{c^n}\, d^{c^{n+1}} \ldots\, d^{c^{n+r-1}} = d^{c^n + c^{n+1} + \ldots + c^{n+r-1}} \geq d^{\log_d(\varepsilon)} = \varepsilon.$$

Then $(x_n)_{n\in\mathbb{N}}$ is forward Cauchy. From the hypothesis of the completeness of (M, R), it follows that there is a limit ℓ of the sequence $(x_n)_{n\in\mathbb{N}}$. Since f is continuous, $f(\ell) = \lim f(x_n) = \lim x_{n+1} = \ell$, and then ℓ is a fixed point for f. Now let us suppose that there is another fixed point ℓ_1. Then

$$R(\ell, \ell_1) = R(f(\ell), f(\ell_1)) \geq (R(\ell, \ell_1))^{c},$$

hence necessarily $R(\ell, \ell_1) = 1$. In the same way we have $R(\ell_1, \ell) = 1$, so from antisymmetry it follows $\ell = \ell_1$.

□

Let us observe that, in particular, the previous theorem holds for the t-norm of the minimum, which is the greatest t-norm, as it is possible to prove immediately.

4.8 Examples of fuzzy orders

In order to give some applications of the notions we discussed above, let us consider some particular fuzzy relations.

Given a set M, we can extend the set theoretical inclusion by means of fuzzy orders defined in a power set $P(M)$. We define a *generalized* $*-$*fuzzy inclusion* as a $*-$fuzzy preorder, i.e. as a reflexive, $*-$transitive fuzzy relation $Incl$: $P(M)\times P(M)\to[0,1]$, such that $X \subseteq Y \Rightarrow Incl(X, Y) = 1$. $Incl$ gives the degree of inclusion of X in Y. If $Incl$ is a generalized $*-$fuzzy inclusion, we have that:

a) $X_1 \subseteq X_2 \Rightarrow Incl(X_1, Y) \geq Incl(X_2, Y)$;

b) $Y_1 \subseteq Y_2 \Rightarrow Incl(X, Y_1) \leq Incl(X, Y_2)$.

Let us observe that a) follows trivially from the $*-$ transitivity:

$$Incl(X_1, Y) \geq Incl(X_1, X_2) * Incl(X_2, Y) \geq 1 * Incl(X_2, Y) = Incl(X_2, Y),$$

and analogously it is possible to prove b).

We say that $Incl$ is a $*-$*fuzzy inclusion* if $Incl$ is a $*-$fuzzy order.

Now let us provide some examples of fuzzy inclusion.

Given a set M, let us define the *fuzzy subset of the relevant element* rel: $M \to [0, 1]$. It is possible to associate to such a fuzzy subset a map μ: $P(M) \to [0, 1]$, defined, for every $X \neq \emptyset$ by

$$\mu(X) = Sup\{rel(x) \,/\, x \in X\}, \tag{4.1}$$

and

$$\mu(\emptyset) = 0.$$

We call such a map the *possibility measure defined by rel*.

We interpret $rel(x)$ as the "degree of relevancy" of an element x and, since it is possible to interpret the existential quantifier in [0,1] by the least upper bound, we interpret $\mu(X)$ as a measure of the truth degree of the claim "there is a relevant element in X".

Proposition 4.8.1 *Let* μ *be the possibility measure defined in* (4.1) *and let us define the fuzzy relation Incl:* $P(M)\times P(M) \to [0, 1]$, *by setting*

$$Incl(X, Y) = 1 - \mu(X - Y). \tag{4.2}$$

Then, Incl is a generalized fuzzy inclusion. Besides, if $rel(x) \neq 0$ *for any* $x \in M$, *then Incl is a fuzzy inclusion.*

Proof. First let us observe that if $X \subseteq Y$, then $X\text{-}Y = \emptyset$, and therefore $Incl(X, Y) = 1$. This proves that *Incl* is an extension of $\subseteq$. Now let us consider $d(X,Y) = \mu\,(X\text{-}Y)$. Proposition 1.8.1 (in Chapter Chapter 1) assures us that proving that *Incl* is a fuzzy preorder is equivalent to prove that d is a generalized quasi-ultrametric. So, we have to prove axioms (d1), (d'1), (d'3) (see Section 1.6) characterizing a generalized quasi-ultrametric. Reflexivity follows trivially by the definition. To prove that $\mu(X\text{-}Z) \leq \mu(X\text{-}Y) \vee \mu(Y\text{-}Z)$, let us observe that

$$X\text{-}Z \subseteq ((X\text{-}Y)\cup(Y\text{-} Z)). \tag{4.3}$$

In fact, let $x \in X\text{-} Z$. If $x\in Y$, then we have that $x\in Y\text{-} Z$, otherwise, if $x\notin Y$, we have that $x\in X\text{-}Y$. Therefore, thanks to (4.3) we can write

$$Sup\{rel(x) \,/\, x\in X\text{-}Z\} \leq Sup\{rel(x) \,/\, x\in (X\text{-}Y)\cup(Y\text{-} Z)\}$$

$$= Sup\{rel(x) \,/\, x\in X\text{-}Y\} \vee Sup\{rel(x) \,/\, x\in Y\text{-} Z\},$$

so the triangular inequality is satisfied and *Incl* is a fuzzy preorder.

To prove the remainder of the proposition, let us assume that $rel(x) \neq 0$ for every $x \in M$. Then $\mu(X) = 0$ entails that $X = \emptyset$. Thus, from $\mu(X\text{-}Y) = 0$ it follows that $X\text{-}Y = \emptyset$, and therefore $X \subseteq Y$. Similarly, from $\mu(Y\text{-}X) = 0$ it follows that $Y\text{-}X = \emptyset$ and $Y \subseteq X$. So, in such a case $d(X, Y) = \mu\,(X\text{-}Y)$ satisfies also (d4) and so *Incl* results a fuzzy inclusion.

□

In accordance with the interpretation of *rel* and μ, we interpret $Incl(X,Y)$ as the truth degree of the claim "there is no relevant element belonging in X and not in Y", or, in other words, "all the relevant elements of X are in Y".

Let us provide a different way to define such a fuzzy inclusion, utilizing the notion of λ-cuts of fuzzy subsets. Given $\lambda \in [0, 1]$, we call *λ-relevant* any element $x \in M$ such that $rel(x) \geq \lambda$ and we identify the set of all λ-relevant elements with the λ-cut $M_\lambda = C(rel, \lambda)$ of the fuzzy subset *rel*. Following this interpretation, a condition like $X \cap M_\lambda \subseteq Y$ means that "every λ-relevant element of X belongs to Y". Let us define $d: P(M) \times P(M) \rightarrow [0, 1]$ by setting

$$d(X, Y) = Inf\{\lambda \in [0,1] \,/\, X \cap M_\lambda \subseteq Y\}. \tag{4.4}$$

Proposition 4.8.2 *Let Incl:* $P(M) \times P(M) \rightarrow [0, 1]$ *be defined by*

$$Incl(X, Y) = 1 - d(X, Y).$$

Then Incl coincides with the generalized fuzzy inclusion defined by (4.2).

Proof. Obviously, from the condition $X \cap M_\lambda \subseteq Y$, if we have $\mu \geq \lambda$ it follows $X \cap M_\mu \subseteq Y$. This means that $\{\lambda \in [0, 1] / X \cap M_\lambda \subseteq Y\}$ is an interval. Thus,

$Inf\{\lambda \in [0, 1] \,/\, X \cap M_\lambda \subseteq Y\} =$

$= Sup\{\lambda \in [0, 1] / X \cap M_\lambda$ is not contained in $Y\}$

$= Sup\{\lambda \in [0, 1] \,/\, x \in X$ exists such that $x \in M_\lambda$ and $x \notin Y\}$

$= Sup\{\lambda \in [0, 1] / x \in X\text{-}Y$ exists such that $rel(x) \geq \lambda\}$

$= Sup\{rel(x) \,/\, x \in X\text{-}Y\} = \mu(X\text{-}Y)$.

□

The just defined class of fuzzy inclusions extends the class given by Seda in the framework of logic programming (see Definition 4.6.3), in a sense. Namely, let n: $M \to N$ be any map and let us set, for every subset X of M,

$$I(X, \lambda) = \{x \in X \,/\, n(x) \leq \lambda\}.$$

Then, we can write the distance defined by Seda in an equivalent way, as the map d': $P(M) \times P(M) \to [0, 1]$ such that

$$d'(X, Y) = Inf\,\{2^{-\lambda} \,/\, I(X, \lambda) \subseteq I(Y, \lambda)\}. \qquad (4.5)$$

We saw that this distance results to be a quasi-ultrametric.
If we choose in a suitable way the map rel and we define a distance d as in (4.4), we prove that such a distance coincides with that defined by Seda. Indeed, we have the following

Proposition 4.8.3 *Let us consider the fuzzy set rel:* $M \to [0, 1]$ *defined as*

$$rel(x) = 2^{-n(x)},$$

and let d and d' be the maps defined in (4.4) *and* (4.5)*, respectively. Then d = d'.*

Proof. Let us observe that

$d(X, Y) = Inf\,\{\lambda \in [0, 1] \,/\, \{x \,/\, 2^{-n(x)} \geq \lambda\} \cap X \subseteq Y\}$

$= Inf\,\{\lambda \in [0, 1] \,/\, \{x \,/\, log_2 2^{-n(x)} \geq log_2(\lambda)\} \cap X \subseteq Y\}$

$= Inf\,\{\lambda \in [0, 1] \,/\, \{x \,/\, n(x) \leq -log_2(\lambda)\} \cap X \subseteq Y\}$

$= Inf\,\{2^{-\lambda} \,/\, \{x \,/\, n(x) \leq \lambda\} \cap X \subseteq Y\} = d'(X, Y)$.

□

Another class of $\ast$-fuzzy inclusions is obtained by assuming that rel: $M \to [0,1]$ satisfies $\sum_{x \in M} rel(x) = 1$. Then, we can define the *finitely additive probability with density* "*rel*" i.e. the map η: $P(M) \to [0,1]$ such that $\eta(\varnothing) = 0$ and, if $X \neq \varnothing$,

$$\eta(X) = \sum_{x \in X} rel(x). \tag{4.6}$$

Differently from $\mu(X)$, $\eta(X)$ takes in account the number of relevant elements in X and therefore we can interpret $\eta(X)$ as a measure of the truth degree of the claim "*there are several relevant elements in X*".

Proposition 4.8.4 *Let f be an additive generator and set*

$$Incl(X, Y) = f^{[-1]}(\eta(X - Y)). \tag{4.7}$$

Then Incl is a generalized $*-$*fuzzy inclusion with respect to t-norm generated by f. If* $rel(x) \neq 0$ *for any* $x \in M$*, then Incl is a* $*-$*fuzzy inclusion.*

Proof. By Proposition 1.8.2, if we prove that $d(X, Y) = \eta(X-Y)$ is a generalized quasi-metrics, it follows that *Incl* is a $*-$fuzzy preorder. Reflexivity follows from the definition. To prove (d3) (see Section 1.6), i.e. that $\eta(X-Z) \leq \eta(X-Y) + \eta(Y-Z)$, let us recall the relation (4.3). Since η is a measure and $(X-Y) \cap (Y-Z) = \emptyset$, we have

$$\eta(X-Z) \leq \eta((X-Y) \cup (Y-Z)) = \eta(X-Y) + \eta(Y-Z).$$

So *Incl* results a $*$ generalised fuzzy inclusion.

To conclude the proof, let us observe that if $rel(x) \neq 0$ for any $x \in M$, then $\eta(X) = 0$ entails that $X = \emptyset$ for any $X \subseteq M$. So from $\eta(X-Y) = 0$ we have that $X-Y = \emptyset$, and therefore $X \subseteq Y$. Likewise, from $\eta(Y-X) = 0$ we have that $Y-X = \emptyset$ and so $Y \subseteq X$. Therefore $d(X, Y) = \eta(X-Y)$ satisfies also (d4) and *Incl* results a $*$-fuzzy inclusion

□

Conclusions and future work

The thesis is devoted to the consideration of two kinds of structures which seem to be useful theoretical tools for information processing. The structures of the first kind are metric in nature and they are obtained by weakening the usual system of axioms for metric spaces in several ways. The structures of the second kind, logical in nature, are strictly related with the interpretation of the equivalences and of the orders in a multi-valued logic (see the notions of similarity, fuzzy order, ...). A duality principle enables us to establish a link between the two classes of notions, i.e. between the metric universe and the logic one.

The researches on the resulting structures are at an initial state and several questions remain open both from a theoretical point of view and with respect to the possible applications. We list some of the issues we will investigate in future work.

On the theoretical side, it is an open question to go on for a point-free approach to geometry in the spirit of Whitehead's ideas. In fact, as sketched in Chapter 2, the notion of *pm*-space is adequate for a point-free approach to the theory of metric spaces. Indeed we can associate any *pm*-space with a metric space and any metric space can be obtained in such a way. Then, it should be interesting searching for a suitable system of axioms to add to the theory of *pm*-spaces, in order to characterize, for instance, the *Euclidean three dimensional pm*-spaces, i.e. the *pm*-spaces whose associated metric space is (isometric with) the three dimensional Euclidean metric space.

A further field of investigation is related with the quoted duality between notions in a metric setting and notions in a logic one. Indeed this duality gives interesting suggestions both in the metric side and in the logic one. For example, in multi-valued logic there is a long time interest for valuation structures different from the ones based on the interval [0, 1]. In accordance, the logical notions of similarity, fuzzy order and so on can be defined in a larger class of structures (for example the class of residuated lattices). So the duality suggests to extend in a suitable way the notion of metric space by admitting distances with values in structures different from the usual set of positive reals. On the other hand the notions of approximate distances and semimetric interval spaces should be dualized into an interesting notion of interval-valued similarity.

With respect to the possible applications of the proposed structures, we are persuaded that whenever the notion of metric space is used in information science, it looks natural to try for an application of the *pm*-spaces when the available information is not complete. For example, the notion of distance in the recognition and classification processes presupposes the identification of the available images with points in a metric space. Now this is correct only in the case of complete information about these images. In the case in which only partial pieces of information are available, then it should be better to refer to the regions in a *pm*-space or in an interval-valued metric space.

Another application is related with logic programming and, in particular, with the question of the deduction in the case of non-monotonic logic. Indeed, the fact that we can unify fixed points theory in an ordered set and fixed points theory in a metric space maybe furnishes news tools for the logic programming field.

Finally, further applications of the proposed distances are, at the present moment, being considered. For instance a distance obtained by the symmetrization of a non-symmetric one was utilized as a tool for improving empirical methods for menu clustering and supporting menu designers for automotive systems (see [4]).

References

[1] Bandler W., Kohout L., Fuzzy power sets and fuzzy implication operators, *Fuzzy Sets and Systems*, 4, 13- 30, 1980.

[2] Brink C., *Power structures, Algebra Universalis*, Vol. 30, (1993), pp. 177-216.

[3] Caprani O., Madsen K., Nielsen H. B., *Introduction to Interval Analysis*, DTU, (2002).

[4] Coppola C., Costagliola G., di Martino S., Ferrucci F., Pacelli T., A fuzzy-based distance to improve empirical methods for menu clustering, submitted to *ICEIS 2006 Conference*.

[5] Coppola C., Miranda A., Pacelli T., Interval semimetric spaces and point-free geometry, *Preprint* n. 3 – 2005, Università di Salerno, DMI.

[6] Coppola C., Pacelli T., Approximate distances, pointless geometry and incomplete information, to appear in *Fuzzy Sets and Systems*, Elsevier.

[7] Coppola C., Pacelli T., Interval semimetric spaces for approximate distances, *Proceedings of the EUSFLAT – LFA 2005 Conference*, 2005, Barcelona, Spain, pp. 1263- 1268.

[8] Crisconio C., *Fuzzy logic and fixed poin theory*, PhD Thesis, University of Napoli, Italy, 2002.

[9] Crisconio C., Gerla G., Similarities and Fuzzy Orders in Approximate Reasoning, Zollo G. (Ed), *New Logics for the New Economy*, Edizioni Scientifiche Italiane (2001), pp. 165- 168.

[10] Di Concilio A., Gerla G., Quasi-metric spaces and point-free geometry, to appear in *Math. Struct. In Comp. Science*.

[11] Diamond P., Kloeden P., Metric spaces of fuzzy sets, *Fuzzy Sets and Systems* 100 Supplement (1999) pp. 63-71.

[12] Dubois D., Prade H., *Fuzzy Sets and Systems: Theory and Applications*, New York Academic Press, 1980.

[13] Dugundji J., Granas A., *Fixed Point Theory*, PWN-Polish Scientific Publishers, 1982.

[14] Düntsh I., A Logic for Rough Sets, *Theoretical Computer Science*, 179(1997a), pp. 427-436.

[15] Düntsh I., Gediga G., Rough set data analysis, *Encyclopedia of Computer Science and Technology*, 2000, 43 (Supplement 28): pp. 281-301.

[16] Fitting M., Fixpoint semantics for logic programming a survey, *Theoretical Computer Science*, 278, (2002), pp. 25-51.

[17] Fitting M., Metric Methods, Three Examples and a Theorem, *J. Logic Programming*, 21 (1994), pp. 113-127.

[18] Fourman M., Scott D.S., *Sheaves and logic, in "Applications of Sheaves"*, Lecture Notes in Mathematics 753, Springer-Verlag (1979), pp. 302-401.

[19] Gerla B., *Many-valued Logics of Continuous t-norms and Their Functional Representation*, PhD Thesis, University of Milano, Italy, 2001.

[20] Gerla G., Coppola C., Pacelli T., Point-free Ultrametric Spaces and the Category of Fuzzy Subsets, *Proceedings of the Tenth International Conference IPMU 2004*, 1, 2004, pp.503-510.

[21] Gerla G., *Fuzzy Logic: Mathematical Tools for Approximate reasoning*, Kluwer Editor.

[22] Gerla G., Miranda A., Graded inclusion and point-free geometry, *Int. J. of Applied Mathematics*, Vol. 11, No 1, 2004, pp. 63-81.

[23] Gerla G., Pointless Geometries, in *Handbook of Incidence Geometries*, F. Buekenhouted (1994) Elsevier Science, pp. 1015-1031.

[24] Gerla G., Pointless Metric Spaces, *Journal of Symbolic Logic* 55, 1, (1990), pp. 207-219.

[25] Gerla G., Representation Theorems for Fuzzy Orders and Quasi-metrics, *Soft Computing*, 8, 2004, pp. 571-580.

[26] Hájek P., *Metamathematics of Fuzzy Logic*, Kluwer Academic Publishers, Dordrecht, 1998.

[27] Hayes B., A Lucid Interval, *American Scientist*, 91, 6 (2003), pp. 484-488.

[28] Höhle U., Presheaves over GL-monoid. In *Nonclassical Logics and Their Applications to Fuzzy Subsets* (U. Höhle, E. P. Klement, eds), Kluwer, (1995), pp. 127-157.

[29] Holly J.E., Pictures of ultrametric spaces, the p-adic numbers, and valued fields, *American Mathematical Monthly* 108 (2001), pp. 721-728.

[30] Liu W.-N., Yao J.T., Yao Y.Y., Rough Approximations under Level Fuzzy Sets, *Lecture Notes in Computer Science, Rough Sets and Current Trends in Computing: 4th International Conference,* 2004. Proceedings, pp. 78-83.

[31] Lloyd J.W., *Foundations of Logic Programming*, Springer-Verlag, 1984.

[32] Menger K., Statistical metrics, *Proc. Nat. Acad. Sci.* 28, 1942, pp. 535-537.

[33] Miyamoto S., *Fuzzy Sets in Information Retrieval and Cluster Analysis*, Kluwer Academic Publishers, Dordrecht, 1990.

[34] Moore R. E., *Interval Analysis*, Englewood Cliffs, New Jersey: Prentice-Hall, 1996.

[35] Novak V., Perfilieva I., Mockor J., *Mathematical Principles of Fuzzy Logic*, Kluwer Academic Publishers, London, 1999.

[36] Pawlak Z., Rough sets, *International Journal of Information and Computer Science*, 11, 1982, pp. 341-356.

[37] Priess-Crampe S., Ribenboim P., Logic Programming and Ultrametric Spaces, *Rend. Mat.*, Serie VII, 19, Roma (1999), pp. 155-176.

[38] Radecki T., Level- fuzzy sets, *J. Cybern.* , 7, 1977, pp. 189-198.

[39] Rutten J.J.M.M., Elements of Generalized Ultrametric Domain Theory, *Technical Report* CS-R9507, Centrum voor Wiskunde en Informatica, Amsterdam, (1995). [Available via FTP at ftp.cwi.nl in directory pub/CWIreports/AP/CS-R9507.ps.gz].

[40] Scarpati L., Gerla G., Extension principles for fuzzy set theory, *Journal of Information Sciences,* 106, (1998), pp. 49-69.

[41] Schweizer B., Sklar A., *Probabilistic metric spaces*, New York: North Holland, 1983.

[42] Seda A. K., Quasi-metrics and Fixed Points in Computing, *Bulletin of the European Association for Theoretical Computer Science* 60 (1996), pp. 154-163.

[43] Seda A. K., Quasi-metrics and the Semantics of Logic Programs, *Fundamenta Informaticae* 29 (2) (1997), pp. 97-117.

[44] Takata O., Miyamoto S., Umayahara K., Clustering of Data with Uncertainties using Hausdorff Distance, *Proc. of the 2nd IEEE International Conference on Intelligent Processing Systems*, Gold Coast, Australia, 1998, pp. 67-71.

[45] Valverde L., On the Structure of F-Indistinguishability Operators, *Fuzzy Sets and Systems*, 17 (1985), pp. 313-328.

[46] Whitehead A. N., *An inquiry concerning the Principles of Natural Knowledge*, Univ. Press, Cambridge, 1919.

[47] Whitehead A. N., *Process and Reality*, McMillian, N.Y., 1929.

[48] Whitehead A. N., *The Concept of Nature*, Univ. Press, Cambridge, 1920.

[49] Yao Y. Y., Two views of the theory of rough sets in finite universes, *International Journal of Approximation Reasoning*, Vol. 15, (1996), No. 4, pp. 291-317.

[50] Yao Y. Y., Wong S. K. M., Interval approaches for uncertain reasoning, Foundations of Intelligence Systems, *Proceedings of the 10th International Symposium on Methodologies of Intelligent Systems*, Lecture Notes in Artificial Intelligence 1325, (1997), pp. 381-390.

[51] Zadeh L. A., Similarity relations and fuzzy orderings, *Inf. Sci.*, 3, pp. 177-200, (1971).

[52] Zadeh L.A., Fuzzy Sets, *Information and Control*, 8, 1965, pp. 338-353.