ALGEBRAIC STRUCTURES FOR ROUGH SETS

GIANPIERO CATTANEO AND DAVIDE CIUCCI

ABSTRACT. Using as example an incomplete information system with support a set of objects X, we discuss a possible algebrization of the concrete algebra of the power set of X through quasi BZ lattices. This structure enables us to define two rough approximations based on a similarity and on a preclusive relation, with the second one always better that the former. Then, we turn our attention to Pawlak rough sets and consider some of their possible algebraic structures. Finally, we will see that also Fuzzy Sets are a model of the same algebras. Particular attention is given to HW algebra which is a strong and rich structure able to characterize both rough sets and fuzzy sets.

1. INTRODUCTION

Rough structures describe the behavior of *concepts*, *properties*, *data*, *abstract objects* in general, that may present some intrinsic vague, ambiguous, unsharp features. For the sake of simplicity we will simply speak of *vague objects*. To these objects it is usually associated some data which characterize them and which can be used to classify them. Data that are often organized as a table associating to each object some attributes values. As observed by Pawlak [Paw98]:

> Data are often presented as a table, columns of which are labelled by *attributes*, rows by *objects* of interest and entries of the table are *attributes values*.
>
> For example, in a table containing information about patients suffering from a certain disease objects are *patients* (strictly speaking their ID's), attributes can be for example *blood pressure*, *body temperature* etc., whereas the entry corresponding to object Smiths and the attribute *blood pressure* can be *normal*.
>
> Such tables are known as *information systems*, *attribute-value tables* or *information tables*.

A natural abstract definition, that reflects this behavior, is the following one.

Definition 1.1. An *Information System* is a structure

$$\mathcal{K}(X) = \langle X, \, Att(X), \, val(X), \, F \rangle$$

where:

- X (called the *universe*) is a non empty set of *objects* (*situations*, *entities*, *states*);
- $Att(X)$ is a non empty set of *attributes*, which assume values for the objects belonging to the set X;
- $val(X)$ is the set of all *possible values* that can be observed for an attribute a from $Att(X)$ in the case of an object x from X;

- F (called the *information map*) is a mapping $F : X \times Att(X) \to val(X)$ which associates to any pair, consisting of an object $x \in X$ and of an attribute $a \in Att(X)$, the value $F(x,a) \in val(X)$ assumed by a for the object x.

In general we assume that an information system satisfies the following two conditions of coherence:

(1) F must be surjective; this means that if there exists a value $v \in val(X)$ which is not the result of the application of the information map F to any pair $(x,a) \in X \times Att(X)$, then this value has no interest with respect to the knowledge stored in the information system.

(2) For any attribute $a \in Att(X)$ there exist at least two objects x_1 and x_2 such that $F(x_1, a) \neq F(x_2, a)$, otherwise this attribute does not supply any knowledge and can be suppressed.

Example 1.2. Imagine you want to rent a flat, and you start to collect information about some apartments. The features you are interested in are: the price of the flat; its location, i.e., if it is down-town or not; the number of rooms and if it has furniture or not. But you are not interested in, for example, at which floor it is located. So, when organizing the data in your possession, you will consider just the first four attributes and omit the floor number attribute. The result is a situation similar to the one presented in Table 1, where the set of objects is $X = \{f_1, f_2, f_3, f_4, f_5, f_6\}$, the family of attributes is $Att(X)=\{$Price, Rooms, Down-Town, Furniture$\}$ and the set of all possible values is $val(X)=\{$high, low, medium, 1, 2, yes, no$\}$.

Flat	Price	Rooms	Down-Town	Furniture
f_1	high	2	yes	no
f_2	high	1	yes	no
f_3	high	2	yes	no
f_4	low	1	no	no
f_5	low	1	no	no
f_6	medium	1	yes	yes

TABLE 1. Flats information system.

■

A generalization of such a concept are *incomplete information systems*, i.e., information systems in which not all values are available for all objects. Such a generalization is justified by empirical observations. Quoting [ST99]:

> An explicit hypothesis done in the classic rough set theory is that all available objects are completely described by the set of available attributes. [...]
> Such a hypothesis, although sound, contrasts with several empirical situations where the information concerning the set A is only partial either because it has not been possible to obtain the attribute values or because it is definitely impossible to get a value for some object on a given attribute.

So, let $\mathcal{K}(X) = \langle X, Att(X), Val(X), F \rangle$ be an incomplete information system. In order to denote the fact that the value possessed by an object x_i with respect to the attribute a_j is unknown, we introduce a *null value* and write $F(x_i, a_j) = *$.

Example 1.3. As a concrete example, let us consider the information system described in Table 1. It can happen that given a flat we do not know all its features, for instance because some information was missing on the advertisement. The result is some missing values in the information system, as shown in Table 2.

Flat	Price	Rooms	Down-Town	Furniture
f_1	high	2	yes	*
f_2	high	*	yes	no
f_3	*	2	yes	no
f_4	low	*	no	no
f_5	low	1	*	no
f_6	*	1	yes	*

TABLE 2. Flats incomplete information system.

∎

Given an information system, complete or not, if we consider pairs of objects belonging to the universe X, we can describe their relationship through a binary relation $\mathcal{R}$. A classification and logical–algebraic characterization of such binary relations can be found in literature (for an overview see [Orl98b]). Generally, these relations are divided into two groups: *indistinguishability* and *distinguishability* relations. In our analysis, we are dealing with a *tolerance* (or *similarity*) relation, i.e., a reflexive and symmetric relation, and its opposite, a *preclusivity* relation, i.e., an irreflexive and symmetric relation. From the intuitive point of view, two individuals are similar when they have an "indistinguishable role" with respect to the intended application, even if they are not equivalent. On the other side, they are preclusive if they have a "distinguishable role", even if they are not totally different.

These considerations lead to the definition of similarity space.

Definition 1.4. A *similarity space* is a structure $\mathcal{S} = \langle X, \mathcal{R} \rangle$, where X (called the *universe* of the space) is a non empty set of objects and $\mathcal{R}$ (called the *similarity* relation of the space) is a reflexive and symmetric binary relation defined on X. In other words:

(i) $\forall x \in X : x\mathcal{R}x$ (reflexivity);
(ii) $\forall x, y \in X : x\mathcal{R}y$ implies $y\mathcal{R}x$ (symmetry).

Let us stress that, in general, a similarity relation is not required to be transitive. This feature strongly differentiate its behavior with respect to the equivalence relation, which is at the basis of classical rough sets theory.

In the context of an incomplete information system $\mathcal{K}(X)$, for a fixed set of attributes $D \subseteq Att(X)$ a natural similarity relation is that two objects are similar if they possess the same values with respect to all known attributes inside D. In a

more formal way:

$$(1.1) \quad \forall x, y \in X: \quad x\mathcal{R}_D y \quad \text{iff} \quad \forall a_i \in D \subseteq Att(X),$$
$$\text{either} \quad F(x, a_i) = F(y, a_i) \quad \text{or} \quad F(x, a_i) = * \quad \text{or} \quad F(y, a_i) = *$$

This is the approach introduced by Kryszkiewicz in [Kry98] which has the advantage that the possibility of null values "corresponds to the idea that such values are just missing, but they do exist. In other words, it is our imperfect knowledge that obliges us to work with a partial information table"[ST99].

Example 1.5. Considering the information system in Table 2, we have, for example, that f_4 is similar to f_5 whatever be the chosen subset $\mathcal{D}$ of attributes. The same applies to f_5 and f_6. On the other hand, the pair of objects f_4 and f_6 are not similar relatively to any subset of attributes which contains the attribute Down–Town; indeed $F(Down - Town, f_4) = no$ whereas $F(Down - Town, f_6) = yes$. In this way, it is also verified that the similarity relation induced from any set of attributes $\mathcal{D}$ which contains Down–Town is not transitive since f_4 is similar to f_5 and f_5 is similar to f_6, but f_4 is not similar to f_6 relatively to $\mathcal{D}$. ∎

Given a similarity space $\langle X, \mathcal{R} \rangle$, the *similarity class* generated by the element $x \in X$ is the collection of all objects similar to x, i.e.,

$$S(x) := \{y \in X : x\mathcal{R}y\}$$

Thus, the similarity class generated by x consists of all the elements which are indistinguishable from x with respect to the similarity relation $\mathcal{R}$. In this way this class constitute a *granule of similarity knowledge* about x and is also called the *granule* generated by x.

Trivially, being $\mathcal{R}$ a reflexive relation, it holds that $x \in S(x)$ and so no similarity class is empty. Further, as a consequence of the non–transitivity of the relation, the similarity classes are not necessarily disjoint, i.e., there may exist x, y such that $S(x) \neq S(y)$ and $S(x) \cap S(y) \neq \emptyset$.

Example 1.6. Making reference to example 1.3, if one considers the set of all attributes (i.e., $D = Att(X)$) and the induced similarity relation according to (1) we have that $S(f_4) = \{f_4, f_5\}$ and $S(f_6) = \{f_2, f_5, f_6\}$. Thus, $S(f_4) \neq S(f_6)$ and $S(f_4) \cap S(f_6) = \{f_5\}$. ∎

Using this notion of similarity class, it is possible to define in a natural way a rough approximation by similarity of any set of objects ([Vak91, SS96, Ste98, ST01]).

Definition 1.7. Given a similarity space $\langle X, \mathcal{R} \rangle$, and a set of objects $A \subseteq X$, the *rough approximation of A by similarity* is defined as the pair $\langle \mathbf{L}_\mathcal{R}(A), \mathbf{U}_\mathcal{R}(A) \rangle$, where

$$(1.2a) \qquad \mathbf{L}_\mathcal{R}(A) := \{x \in X : S(x) \subseteq A\} = \{x \in X : \forall z \ (x\mathcal{R}z \Rightarrow z \in A)\}$$

$$(1.2b) \qquad \mathbf{U}_\mathcal{R}(A) := \{x \in X : S(x) \cap A \neq \emptyset\} = \{x \in X : \exists z \ (x\mathcal{R}z \text{ and } z \in A)\}$$

It is easy to verify that the following chain of inclusions holds:

$$(1.3) \qquad\qquad\qquad \mathbf{L}_\mathcal{R}(A) \subseteq A \subseteq \mathbf{U}_\mathcal{R}(A).$$

$\mathbf{L}_\mathcal{R}(A)$ is called the *similarity lower* approximation of A (since it is constituted by all objects whose granule is contained in A) and $\mathbf{U}_\mathcal{R}(A)$ the *similarity upper*

approximation of A (since it is constituted by all objects whose granule has at least a point in common with A).

Example 1.8. Consider again the information system in Table 2 with similarity relation (1) relative to $D = Att(X)$. Choosing the set of objects $A = \{f_4, f_5, f_6\}$, we have $\mathbf{L}_\mathcal{R}(A) = \{f_4, f_5\}$ and $\mathbf{U}_\mathcal{R}(A) = \{f_2, f_4, f_5, f_6\}$ with the obvious satisfaction of the chain (3). ∎

As said before, the opposite of a similarity relation is a preclusive relation: two objects are in a preclusive relation iff it is possible to distinguish one from the other. Using such a relation it is possible to define a notion dual to the one of similarity space.

Definition 1.9. A *preclusivity space* is a structure $\mathcal{S} = \langle X, \# \rangle$, where X (called the *universe* of the space) is a non empty set and $\#$ (called the *preclusivity* relation of the space) is an irreflexive and symmetric relation defined on X. In other words:

   (i) $\forall x \in X :$ not $x \# x$ (irreflexivity);
   (ii) $\forall x, y \in X : x \# y$ implies $y \# x$ (symmetry).

Needless to stress, any similarity space $\langle X, \mathcal{R} \rangle$ determines a corresponding preclusivity space $\langle X, \#_\mathcal{R} \rangle$ with $x \#_\mathcal{R} y$ iff $\neg(x\mathcal{R}y)$, and vice versa any preclusivity space $\langle X, \# \rangle$ determines a similarity space $\langle X, \mathcal{R}_\# \rangle$ with $x\mathcal{R}_\# y$ iff $\neg(x\#y)$. In this case we will say that we have a pair of *correlated* similarity–preclusive relations.

Suppose, now, a preclusivity space $\langle X, \# \rangle$. The preclusive relation $\#$ permits us to introduce for all $H \in \mathcal{P}(X)$ (where we denote by $\mathcal{P}(X)$ the power set of X) its *preclusive complement* defined as

$$H^\# := \{x \in X : \forall y \in H \ (x \# y)\}.$$

In other words, $H^\#$ contains all and only the elements of X that are distinguishable from all the elements of H. Whenever $x \in H^\#$ we will also write: $x \# H$ and we will say that x is *preclusive* to the set H. Further, we will say that two subsets H and K of X are *mutually preclusive* ($H \# K$) iff all the elements of H are preclusive to K, and all the elements of K are preclusive to H. We remark that, in the context of modal analysis of rough approximation spaces, the operation $\#$ is a *sufficiency* operator [DO01].

The following result has been proved in [Cat97].

Proposition 1.10. *Let* $\mathcal{S} = \langle X, \# \rangle$ *be a preclusivity space and let us consider the structure* $\langle \mathcal{P}(X), \cap, \cup, {}^c, \#, \emptyset, X \rangle$. *Then, the following hold:*

   (1) *The substructure* $\langle \mathcal{P}(X), \cap, \cup, \emptyset, X \rangle$ *is a distributive (complete) lattice, with respect to standard set theoretic intersection* $\cap$ *and union* $\cup$, *bounded by the least element* $\emptyset$ *and the greatest element* X. *The partial order relation induced from this lattice structure is the standard set theoretic inclusion* $\subseteq$.
   (2) *The operation* ${}^c : \mathcal{P}(X) \to \mathcal{P}(X)$ *associating to any subset* H *of* X *its set theoretic complement* $H^c = X \setminus H$ *is a standard complementation, i.e., it satisfies:*

   (C-1) $H = (H^c)^c$ *(involution)*

   (C-2a) $H \subseteq K$ *iff* $K^c \subseteq H^c$ *(contraposition)*
   (C-2b) $H^c \cap K^c = (H \cup K)^c$ *($\cap$ de Morgan)*
   (C-2c) $H^c \cup K^c = (H \cap K)^c$ *($\cup$ de Morgan)*

(C-3a) $H \cap H^c = \emptyset$ (non contradiction)
(C-3b) $H \cup H^c = X$ (excluded middle)

(3) *The operator* $^{\#} : \mathcal{P}(X) \to \mathcal{P}(X)$ *associating to any subset* H *of* X *its preclusive complement* $H^{\#}$ *is a unusual complementation and it satisfies:*

(B1) $H \subseteq (H^{\#})^{\#}$ *(weak double negation)*

(B2a) *If* $H \subseteq K$ *then* $K^{\#} \subseteq H^{\#}$ *(contraposition)*
(B2b) $H^{\#} \cap K^{\#} = (H \cup K)^{\#}$ *($\cap$ de Morgan for $\#$)*

(B3) $H \cap H^{\#} = \emptyset$ *(non contradiction)*
(4) *The following interconnection rule holds:*
(in) $H^{\#} \subseteq H^c$.

In the particular case of the similarity relation (1), the corresponding preclusive relation is just its logical negation:

(1.4) $\forall x, y \in X$ $x \#_D y$ iff   not $x \mathcal{R}_D y$ iff

$\exists a_i \in D \subseteq Att(X) : F(x, a_i) \neq F(y, a_i)$ and $F(x, a_i) \neq *$ and $F(y, a_i) \neq *$

Example 1.11. Consider, now, the set of flats $H = \{f_2, f_6\}$ relatively to Table 2, then the preclusive complement of H with respect to $D = Att(X)$ is $H^{\#} = \{f_4\}$, which is different from $H^c = \{f_1, f_3, f_4, f_5\}$. Further, the complementation $^{\#}$ does not satisfy some standard properties. First of all the strong double negation law, $H = H^{\#\#}$, and the excluded middle law $H \cup H^{\#} = X$. In fact, we have

$$H = \{f_2, f_6\} \subsetneq \{f_1, f_2, f_3, f_6\} = H^{\#\#}$$

and

$$H \cup H^{\#} = \{f_2, f_4, f_6\} \neq X$$

Moreover, also the $\cup$ de Morgan law, $H^{\#} \cup K^{\#} = (H \cap K)^{\#}$, and the strong contraposition law, $H^{\#} \subseteq K^{\#}$ implies $K \subseteq H$, are not generally verified. Let us consider, for instance, $K = \{f_1, f_3\}$. Then, we obtain

$$(H \cap K)^{\#} = \emptyset^{\#} = X \neq \{f_4, f_5, f_6\} = H^{\#} \cup K^{\#}$$

and

$$H^{\#} = \{f_4\} \subseteq \{f_4, f_5, f_6\} = K^{\#} \quad \text{but} \quad K = \{f_1, f_3\} \not\subseteq \{f_2, f_6\} = H$$

Finally, in general, it does not hold the strong interconnection rule $H^{\#c} = H^{\#\#}$. For instance, $H^{\#\#} = \{f_1, f_2, f_3, f_6\} \neq \{f_1, f_2, f_3, f_5, f_6\} = H^{\#c}$.   ∎

On the Boolean lattice based on the power set $\mathcal{P}(X)$, we now have two, generally different, complementations: the usual set theoretic complementation ^{c} and the preclusive complementation $^{\#}$. By their interaction, it is possible to define a closure and an interior operator on $\mathcal{P}(X)$.

Proposition 1.12. *Let* $\langle \mathcal{P}(X), \cap, \cup, {}^c, {}^{\#}, \emptyset, X \rangle$ *be the algebraic structure based on the power set of* X *and generated by the preclusivity space* $\langle X, \# \rangle$. *Then the mapping*

$$\mathbf{L}_{\#} : \mathcal{P}(X) \to \mathcal{P}(X), \quad H \to \mathbf{L}_{\#}(H) := H^{c\#\#c}$$

is an interior operator, *i.e.,*

(I_0) $\quad X = \mathbf{L}_\#(X)$ $\hfill$ (*normalized*)

(I_1) $\quad \mathbf{L}_\#(H) \subseteq H$ $\hfill$ (*decreasing*)

(I_2) $\quad \mathbf{L}_\#(H) = \mathbf{L}_\#(\mathbf{L}_\#(H))$ $\hfill$ (*idempotent*)

(I_3) $\quad \mathbf{L}_\#(H \cap K) \subseteq \mathbf{L}_\#(H) \cap \mathbf{L}_\#(K)$ $\hfill$ (*sub − multiplicative*)

Under condition $(I_0) - (I_1)$, the condition (I_3) is equivalent to the *monotonicity condition*: $H \subseteq K$ implies $\mathbf{L}_\#(H) \subseteq \mathbf{L}_\#(K)$. In general, the Kuratowsky property of topological interior, $\mathbf{L}_\#(H \cap K) = \mathbf{L}_\#(H) \cap \mathbf{L}_\#(K)$, is not verified.

Example 1.13. Let us consider the two sets of flats $H = \{f_1, f_2, f_4, f_5\}$ and $K = \{f_1, f_2, f_5, f_6\}$, relative to Table 2. Then, the interior of H, K and their intersection, relative to $D = Att(X)$, are respectively $\mathbf{L}_\#(H) = \{f_4, f_5\}$, $\mathbf{L}_\#(K) = \{f_2, f_5, f_6\}$ and $\mathbf{L}_\#(\{f_1, f_2, f_5\}) = \emptyset$. So, $\mathbf{L}_\#(H \cap K) = \emptyset \neq \{f_5\} = \mathbf{L}_\#(H) \cap \mathbf{L}_\#(K)$. ■

From the fact that according to (I_1) one has that $\mathbf{L}_\#(H) \subseteq H$, it is possible to single out the collection of all #–*open* sets defined as follows:

$$\mathbb{O}(X, \#) := \{A \subseteq X : A = \mathbf{L}_\#(A) = A^{c\#\#c}\}.$$

This set is not empty since both the empty set $\emptyset$ and the whole universe X are #–open sets.

Proposition 1.14. *Let* $\langle \mathcal{P}(X), \cap, \cup, {}^c, \emptyset, X \rangle$ *be the algebraic structure generated by the preclusivity space* $\langle X, \# \rangle$ *and* $^\#$ *the preclusive complement on* $\mathcal{P}(X)$. *Then, the mapping*

$$\mathbf{U}_\# : \mathcal{P}(X) \to \mathcal{P}(X), \quad H \to \mathbf{U}_\#(H) := H^{\#\#}$$

is a closure operator, *i.e.,*

(C_0) $\quad \emptyset = \mathbf{U}_\#(\emptyset)$ $\hfill$ (*normalized*)

(C_1) $\quad H \subseteq \mathbf{U}_\#(H)$ $\hfill$ (*increasing*)

(C_2) $\quad \mathbf{U}_\#(H) = \mathbf{U}_\#(\mathbf{U}_\#(H))$ $\hfill$ (*idempotent*)

(C_3) $\quad \mathbf{U}_\#(H) \cup \mathbf{U}_\#(K) \subseteq \mathbf{U}_\#(H \cup K)$ $\hfill$ (*sub − additive*)

Let us note that condition (C_3), under $(C_0) - (C_1)$, is equivalent to the *monotonicity condition*: $H \subseteq K$ implies $\mathbf{U}_\#(H) \subseteq \mathbf{U}_\#(K)$. In general, the Kuratowski property of topological closure, $\mathbf{U}_\#(H) \cup \mathbf{U}_\#(K) = \mathbf{U}_\#(H \cup K)$, does not hold.

Example 1.15. Let us consider the two sets of flats $H = \{f_2\}$ and $K = \{f_5\}$, relative to Table 2. Then, the closure of H, K and their union, relative to $D = Att(X)$, are respectively $\mathbf{U}_\#(H) = \{f_1, f_2, f_3\}$, $\mathbf{U}_\#(K) = \{f_4, f_5\}$ and $\mathbf{U}_\#(\{f_2, f_5\}) = X$. So, $\mathbf{U}_\#(H \cup K) = X \neq \{f_1, f_2, f_3, f_4, f_5\} = \mathbf{U}_\#(H) \cup \mathbf{U}_\#(K)$. ■

Since according to (C_1) one has that $H \subseteq \mathbf{U}_\#(H)$, it is possible to introduce the collection of all #–*closed* sets defined as follows:

$$\mathbb{C}(X, \#) := \{B \subseteq X : B = \mathbf{U}_\#(B) = B^{\#\#}\}.$$

Also this set is not empty since the empty set $\emptyset$ and the whole universe X are both #–closed. It is easy to see that A is #–open iff A^c is #–closed, and similarly B is

#–closed iff B^c is #–open. Hence,

$$\mathbb{O}(X, \#) = \{A \subseteq X : A^c \in \mathbb{C}(X, \#)\}$$
$$\mathbb{C}(X, \#) = \{B \subseteq X : B^c \in \mathbb{O}(X, \#)\}.$$

If a set is both #–open and #–closed, it is said to be #–*clopen*. So, the collection of all #–clopen sets is defined as:

$$\mathbb{CO}(X, \#) = \mathbb{C}(X, \#) \cap \mathbb{O}(X, \#).$$

Both the empty set $\emptyset$ and the whole universe X are #–clopen. In the sequel, if there is no confusion, we simply speak of open, closed, and clopen sets instead of #–open, #–closed, and #–clopen sets.

By the increasing (resp., decreasing) property (C_1) (resp., (I_1)) of the closure (resp., interior) operator, the following holds:

(1.5) $$\mathbf{L}_\#(H) \subseteq H \subseteq \mathbf{U}_\#(H).$$

Therefore, the pair $\langle \mathbf{L}_\#(H), \mathbf{U}_\#(H) \rangle$ can be thought of as a preclusive rough approximation of the set H by a preclusive open–closed pair. In fact,

   (1) by (C_2) we have that $\mathbf{U}_\#(H)$ is a closed set, and inclusion $H \subseteq \mathbf{U}_\#(H)$ says that $\mathbf{U}_\#(H)$ is one of the possible upper closed approximations of H;
   (2) for every closed set B which is an upper approximation of H, $H \subseteq B$, we have by monotonicity that $\mathbf{U}_\#(H) \subseteq \mathbf{U}_\#(B) = B$,
   (3) thus, $\mathbf{U}_\#(H)$ is the best approximation of H from the top by closed elements (the *upper closed approximation* of H).

By the above properties (1)–(3), it follows that the preclusive upper approximation of a set H can also be expressed as:

(1.6) $$\mathbf{U}_\#(H) = \cap\{B \in \mathbb{C}(X, \#) : H \subseteq B\}.$$

In the case of a closed set H one has that $H = \mathbf{U}_\#(H)$, i.e., the upper closed approximation of any closed set is the set itself. In this sense we can say that closed sets are *upper exact sets*.

Analogously, for the lower preclusive approximation:

   (1) by (I_2) we have that $\mathbf{L}_\#(H)$ is an open set, and inclusion $\mathbf{L}_\#(H) \subseteq H$ says that $\mathbf{L}_\#(H)$ is one of the possible lower open approximations of H;
   (2) for every open set B which is a lower approximation of H, $B \subseteq H$, we have by monotonicity that $B = \mathbf{L}_\#(B) \subseteq \mathbf{L}_\#(H)$,
   (3) thus, $\mathbf{L}_\#(H)$ is the best approximation of H from the bottom by -open elements (the *lower open approximation* of H).

By the above properties (1)–(3), it follows that the preclusive lower approximation of a set H can be expressed as:

(1.7) $$\mathbf{L}_\#(H) = \cup\{B \in \mathbb{C}(X, \#) : B \subseteq H\}.$$

In the case of an open set H obviously $H = \mathbf{L}_\#(H)$, obtaining that open sets can be considered as *lower exact sets*.

Finally, we have that clopen sets are both lower and upper exact sets, so, we simply call them *exact sets*.

Example 1.16. Making reference to the information system described by Table 2, let us consider the set $H = \{f_4, f_5, f_6\}$ and compute its preclusive rough approximation with respect to the preclusive relation (4) induced by the set $\mathcal{D} = Att(X)$

of all involved attributes. We have that $H^{\#} = \{f_1, f_3\}$ and $\mathbf{U}_{\#}(H) = H^{\#\#} = \{f_4, f_5, f_6\} = H$. Computing the preclusive lower approximation, one obtains that $H^c = \{f_1, f_2, f_3\}$, $H^{c\#} = \{f_4, f_5\}$, $H^{c\#\#} = \{f_1, f_2, f_3\}$ and finally, $H^{c\#\#c} = \mathbf{L}_{\#}(H) = \{f_4, f_5, f_6\} = H$. Thus, the rough approximation of $H = \{f_4, f_5, f_6\}$ is the pair $\langle\{f_4, f_5, f_6\}, \{f_4, f_5, f_6\}\rangle$, and H is a clopen set.

In Figures 1–3, the closed, open, and clopen sets relative to the preclusive relation induced from Table 2 by the set $\mathcal{D} = Att(X)$ are represented through Hasse diagrams.
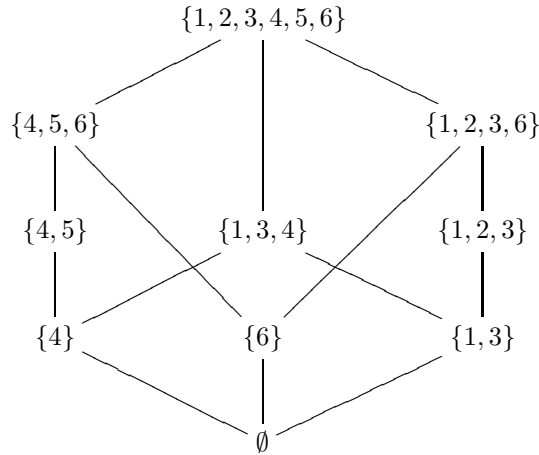

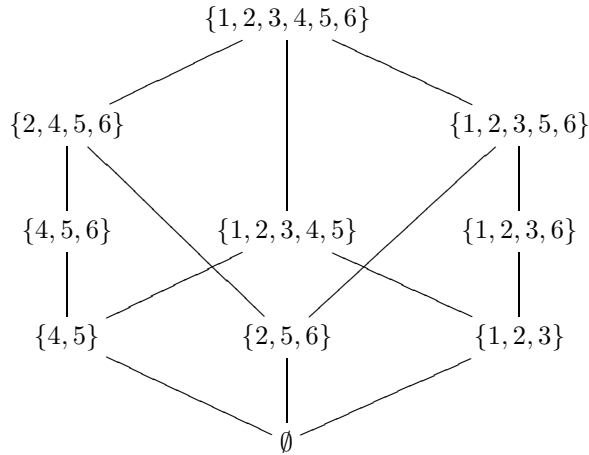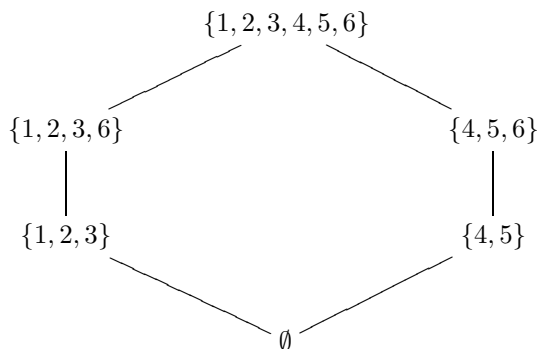FIGURE 1. Closed sets relative to Table 2 for $\mathcal{D} = Att(x)$.

FIGURE 2. Open sets relative to Table 2 for $\mathcal{D} = Att(x)$.

$$\{1,2,3,4,5,6\}$$

$$\{1,2,3,6\} \qquad\qquad \{4,5,6\}$$

$$\{1,2,3\} \qquad\qquad \{4,5\}$$

$$\emptyset$$

FIGURE 3. Clopen sets relative to Table 2 for $\mathcal{D} = Att(x)$.

In Table 3 examples of similar and preclusive approximations with respect to some particular subsets of the involved universe are reported.

$\mathbf{L}_{\mathcal{R}}(H)$	$\mathbf{L}_{\#}(H)$	H	$\mathbf{U}_{\#}(H)$	$\mathbf{U}_{\mathcal{R}}(H)$
$\emptyset$	$\emptyset$	$\{f_2\}$	$\{f_1,f_2,f_3\}$	$\{f_1,f_2,f_3,f_6\}$
$\{f_4\}$	$\{f_4,f_5\}$	$\{f_4,f_5\}$	$\{f_4,f_5\}$	$\{f_4,f_5,f_6\}$
$\{f_4\}$	$\{f_4,f_5\}$	$\{f_1,f_2,f_4,f_5\}$	$\{f_1,f_2,f_3,f_4,f_5,f_6\}$	$\{f_1,f_2,f_3,f_4,f_5,f_6\}$
$\{f_6\}$	$\{f_2,f_5,f_6\}$	$\{f_1,f_2,f_5,f_6\}$	$\{f_1,f_2,f_3,f_4,f_5,f_6\}$	$\{f_1,f_2,f_3,f_4,f_5,f_6\}$

TABLE 3. Examples of Approximations.

As can be seen, in all these particular cases the following chain of inclusions holds:

$$(1.8) \qquad \mathbf{L}_{\mathcal{R}}(H) \subseteq \mathbf{L}_{\#}(H) \subseteq H \subseteq \mathbf{U}_{\#}(H) \subseteq \mathbf{U}_{\mathcal{R}}(H)$$

Thus, it seems possible to conjecture that the preclusive approximation gives a better, i.e., closer, approximation of a given set H than the corresponding similarity one. We will show in Section 3 that this is a general property, which holds for all subsets of the universe of a preclusivity space. ∎

2. About the monotonical increase in time of knowledge

Starting from an incomplete information system one can wonder what happens when we have more knowledge about it, i.e., when the number of unknown values decreases. One could expect that for a fixed set of attributes to more knowledge there corresponds more open and closed sets producing in this way a preclusive rough approximation which is better than the previous one (if the information increases, then the approximation should be better and better). However, this is not always the case as the following examples will show.

Example 2.1. We now consider two examples showing two different cases making reference to the information system of Table 2 which is supposed relative to a given time, say t_1. Let us consider a first situation, which can be attributed to a knowledge situation of some observer O_1 at a time, say t_0, former than the time t_1

(i.e., $t_0 < t_1$) in which the knowledge of Table 2 has been achieved. This first case corresponds to a worst knowledge about flat f_3 according to the situation described in Table 4.

Flat	Price	Rooms	Down-Town	Furniture
f_1	high	2	yes	*
f_2	high	*	yes	no
f_3	*	2	*	no
f_4	low	*	no	no
f_5	low	1	*	no
f_6	*	1	yes	*

TABLE 4. Case 1: only the information regarding flat f_3 is different from the one of Table 2 with a global increase in time of the knowledge.

In the second case, another observer O_2 is involved whose former knowledge, relative to time t_0, with respect to the information system of Table 2 consists of a worst knowledge about flat f_6 according to the situation described in Table 5.

Flat	Price	Rooms	Down-Town	Furniture
f_1	high	2	yes	*
f_2	high	*	yes	no
f_3	*	2	yes	no
f_4	low	*	no	no
f_5	low	1	*	no
f_6	*	1	*	*

TABLE 5. Case 2: only the information regarding flat f_6 is different from the one of Table 2 with a global increase in time of the knowledge.

Thus, there are two observers that initially (time t_0) have a different knowledge about the same collection of flats (described by Tables 4 and 5). During the time, both the observers increase their knowledge reaching the same result exposed in Table 2 at a subsequent time t_1. The final result is that relatively to the same set of all attributes $\mathcal{D} = Att(X)$ there is a decrease in the number of closed (and so also of open) sets passing from case 1 (time t_0) to the situation of Table 2 (time t_1) and an increase going from case 2 (time t_0) to Table 2 (time t_1), as can be seen looking at Figures 4 and 5.

$$\{1,2,3,4,5,6\}$$

$$\{1,3,4\} \qquad \{1,2,3\} \qquad \{1,2,6\} \qquad \{4,5,6\}$$

$$\{1,3\} \qquad \{1,2\} \qquad \{4,5\} \qquad \{5,6\}$$

$$\{1\} \qquad \{4\} \qquad \{5\} \qquad \{6\}$$
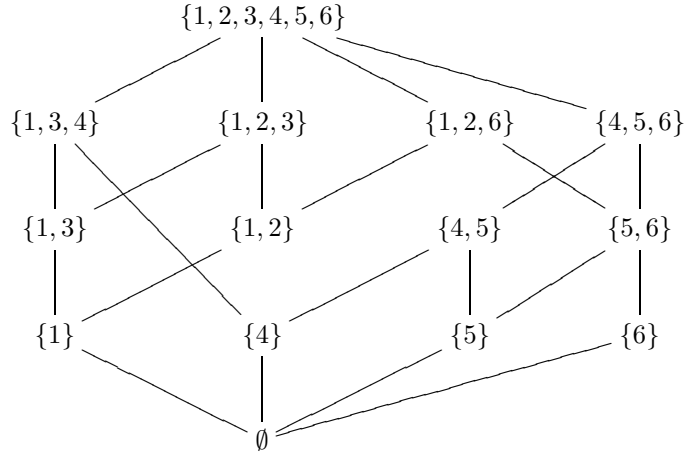
$$\emptyset$$

FIGURE 4. Closed sets in case 1 at time t_0: during the time evolution from t_0 to t_1 (corresponding to Figure 1) there is a decrease of the total number of closed sets with the appearance of the set $\{1,2,3,6\}$.
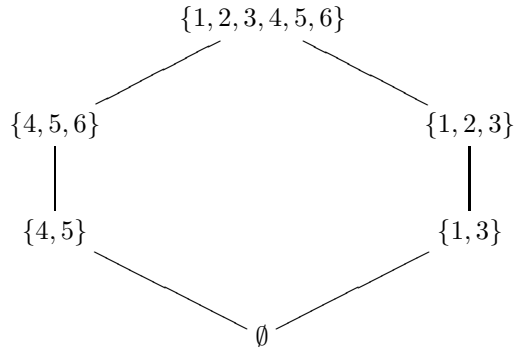
$$\{1,2,3,4,5,6\}$$

$$\{4,5,6\} \qquad \{1,2,3\}$$

$$\{4,5\} \qquad \{1,3\}$$

$$\emptyset$$

FIGURE 5. Closed sets in case 2 at time t_0: during the time evolution from t_0 to t_1 (corresponding to Figure 1) there is an increase of the total number of closed sets.

When considering the clopen sets we observe that their number increases in the original situation depicted in figure 3 with respect to both case 1 and case 2, as can be seen in Figure 6.
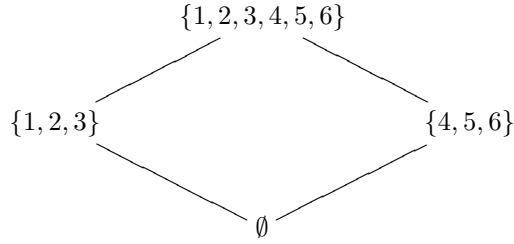
$$\{1, 2, 3, 4, 5, 6\}$$

$$\{1, 2, 3\} \qquad \{4, 5, 6\}$$

$$\emptyset$$

FIGURE 6. Clopen Sets: case 1 and case 2 at time t_0.

Again we ask whether this is a general property: to a greater knowledge corresponds a higher number of clopen sets. Also in this case, the answer is negative. Let us suppose that, with respect to the original situation of Table 2 at time t_1, both the observers in a later time, say t_2, increase their knowledge about flat 5 according to Table 6.

Flat	Price	Rooms	Down-Town	Furniture
f_1	high	2	yes	*
f_2	high	*	yes	no
f_3	*	2	yes	no
f_4	low	*	no	no
f_5	low	1	no	no
f_6	*	1	yes	*

TABLE 6. Case 3 at time $t_2 > t_1$.

In this case, however, the number of clopen sets decreases with respect to the knowledge at time t_1, as can be seen in Figure 7.

$$\{1, 2, 3, 4, 5, 6\}$$

$$\{4, 5\} \qquad \{1, 2, 3, 6\}$$
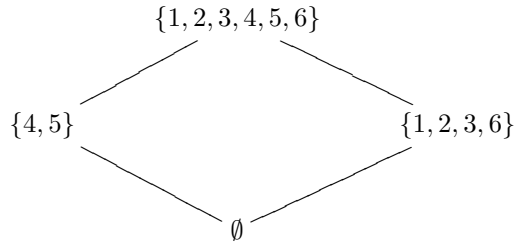
$$\emptyset$$

FIGURE 7. Clopen Sets: Case 3 at time $t_2 > t_1$.

When considering the closed sets, it happens that they are numerically less at time t_2 with respect to the ones at time t_1, but the set $\{1, 3, 4, 5\}$ is closed at time t_2 and not at time t_1.

In Figure 8 it is reported the Hasse diagram of closed sets at time t_2.
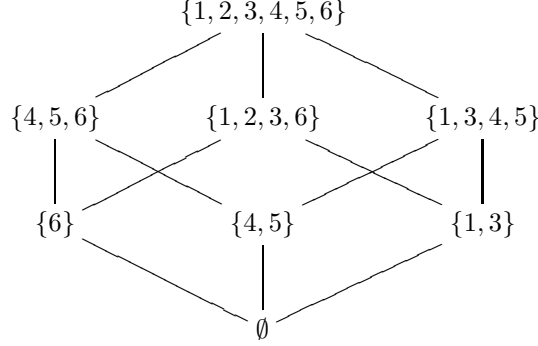

FIGURE 8. Closed sets: Case 3 at time $t_2 > t_1$.

Finally, as a limit final case we consider the complete information system of Table 1, corresponding to an increase of the knowledge represented by Table 6 (occurred at a time $t_3 > t_2$). In this case the set of closed and clopen sets are the same as the ones of case 3 (see Figures 8 and 7 respectively). ■

We can conclude that there is no correspondence between unknown values and exact (either open or closed or clopen sets) of the preclusive environment. The same uncertainty holds with respect to the quality of an approximation. Let us consider an incomplete information system based on the universe X with associated a pair $(\mathcal{R}, \#)$ of correlated similarity–preclusive relations. For a given set of objects H from X we can compute its preclusive rough approximation. Now, if we add (resp., remove) some information to (resp., from) the system, and recompute the preclusive approximation, we have no idea if the result is better or worst than the previous one.

Example 2.2. Let us consider the information system at time t_0 of Table 4. If we compute the approximations of the same sets used in Table 3, we obtain the following results.

$\mathbf{L}_{\mathcal{R}}(H)$	$\mathbf{L}_{\#}(H)$	H	$\mathbf{U}_{\#}(H)$	$\mathbf{U}_{\mathcal{R}}(H)$
$\emptyset$	$\emptyset$	$\{f_2\}$	$\{f_1, f_2\}$	$\{f_1, f_2, f_3, f_6\}$
$\emptyset$	$\emptyset$	$\{f_4, f_5\}$	$\{f_4, f_5\}$	$\{f_3, f_4, f_5, f_6\}$
$\emptyset$	$\emptyset$	$\{f_1, f_2, f_4, f_5\}$	$\{f_1, f_2, f_3, f_4, f_5, f_6\}$	$\{f_1, f_2, f_3, f_4, f_5, f_6\}$
$\{f_6\}$	$\{f_2, f_5, f_6\}$	$\{f_1, f_2, f_5, f_6\}$	$\{f_1, f_2, f_3, f_4, f_5, f_6\}$	$\{f_1, f_2, f_3, f_4, f_5, f_6\}$

TABLE 7. Examples of Approximations, case 1 at time t_0.

That is, going from a situation at time t_0 to a situation at time t_1, i.e., adding knowledge to the information system in a monotonic way, we have that

(1) the time evolution of the preclusive rough approximation of a set is unpredictable, i.e., the approximation becomes either worst (case $\{f_2\}$), or better (cases $\{f_4, f_5\}$, $\{f_1, f_2, f_4, f_5\}$), or remains the same (case $\{f_1, f_2, f_5, f_6\}$);

(2) on the contrary, in this example the similarity rough approximation at time t_1 of a set is always better than the corresponding similarity rough approximation of the same set at the previous time t_0, i.e., there is a *monotone* behavior of the similarity approximation with respect to the knowledge increase;

(3) however, at any fixed time the preclusive approximation of a set is always better then the correlated similarity approximation, i.e., the chain (8) holds for any set.

■

As regards to the preclusive approximations, we can select those situations in which to an increase of knowledge in time corresponds an increase in the quality of the approximations.

First of all, let us introduce the following definitions.

Definition 2.3. Let $\mathcal{K}^{(t_0)}(X)$ and $\mathcal{K}^{(t_1)}(X)$, with $t_0, t_1 \in \mathbb{R}$, be two incomplete information systems based on the same triple $\langle X, Att(X), val(X) \rangle$ and characterized by two different information maps

$$F^{(t_0)} : X \times Att(X) \mapsto val(X) \quad \text{and} \quad F^{(t_1)} : X \times Att(X) \mapsto val(X).$$

We will say that there is a *monotonic increase of information* iff

$$t_0 \leq t_1 \quad \text{and} \quad \forall(x, a) \ (F^{(t_0)}(x, a) \neq * \ \Rightarrow \ F^{(t_1)}(x, a) = F^{(t_0)}(x, a)).$$

In such a case, we will write $\mathcal{K}^{(t_0)}(X) \preceq \mathcal{K}^{(t_1)}(X)$.

Definition 2.4. Let $\mathcal{K}^{(t_0)}(X)$ and $\mathcal{K}^{(t_1)}(X)$ be two incomplete information systems such that $\mathcal{K}^{(t_0)}(X) \preceq \mathcal{K}^{(t_1)}(X)$. We will say that there is a *monotonic increase of knowledge* iff $\mathbb{C}^{(t_0)}(X) \subseteq \mathbb{C}^{(t_1)}(X)$.

In the previous examples we have seen that to a monotonic increase of information does not correspond a monotonic increase in the quality of the approximations. This desirable behavior holds, instead, in the case of monotonic increase of knowledge.

Proposition 2.5. *Let $\mathcal{K}^{(t_0)}(X)$ and $\mathcal{K}^{(t_1)}(X)$ be two incomplete information systems characterized by a monotonic increase of knowledge . Then*

$$(2.1) \qquad \forall H \subseteq X \quad \mathbf{L}_{\#}^{(t_0)}(H) \subseteq \mathbf{L}_{\#}^{(t_1)}(H) \subseteq H \subseteq \mathbf{U}_{\#}^{(t_1)}(H) \subseteq \mathbf{U}_{\#}^{(t_0)}(H).$$

Proof. By hypothesis, $\mathbb{C}^{(t_0)}(X) \subseteq \mathbb{C}^{(t_1)}(X)$ and consequently $\mathbb{O}^{(t_0)}(X) \subseteq \mathbb{O}^{(t_1)}(X)$. Now, let us consider the rough lower approximation of a set H, as defined in equation (7). Suppose that there exists a set $A \in \mathbb{O}^{(t_0)}(X)$ such that $A \subseteq H$, then, by hypothesis, $A \in \mathbb{O}^{(t_1)}(X)$. So $\mathbf{L}_{\#}^{(t_0)}(H) = \cup\{A \in \mathbb{O}^{(t_0)}(X) : A \subseteq H\} \subseteq \cup\{A \in \mathbb{O}^{(t_1)}(X) : A \subseteq H\} = \mathbf{L}_{\#}^{(t_1)}(H)$. On the other hand, suppose that there exists $B \in \mathbb{C}^{(t_0)}(X)$ such that $H \subseteq B$, then it also holds $B \in \mathbb{C}^{(t_1)}(X)$. So, $\mathbf{U}_{\#}^{(t_1)}(H) = \cap\{B \in \mathbb{C}^{(t_1)}(X) : H \subseteq B\} \subseteq \cap\{B \in \mathbb{C}^{(t_0)}(X) : H \subseteq B\} = \mathbf{U}_{\#}^{(t_0)}(H)$ □

Example 2.6. Let us consider the case 2, at time t_0, and compute the approximations of the same sets of Table 3. The results are shown in Table 8.

$\mathbf{L}_{\mathcal{R}}(H)$	$\mathbf{L}_{\#}(H)$	H	$\mathbf{U}_{\#}(H)$	$\mathbf{U}_{\mathcal{R}}(H)$
$\emptyset$	$\emptyset$	$\{f_2\}$	$\{f_1, f_2, f_3\}$	$\{f_1, f_2, f_3, f_6\}$
$\emptyset$	$\emptyset$	$\{f_4, f_5\}$	$\{f_4, f_5\}$	$\{f_4, f_5, f_6\}$
$\emptyset$	$\emptyset$	$\{f_1, f_2, f_4, f_5\}$	$\{f_1, f_2, f_3, f_4, f_5, f_6\}$	$\{f_1, f_2, f_3, f_4, f_5, f_6\}$
$\emptyset$	$\emptyset$	$\{f_1, f_2, f_5, f_6\}$	$\{f_1, f_2, f_3, f_4, f_5, f_6\}$	$\{f_1, f_2, f_3, f_4, f_5, f_6\}$

TABLE 8. Examples of Approximations, case 2.

Going from Table 5 (case 2), at time t_0, to Table 2, at time t_1, we have a monotonic increase of knowledge. So, as expected, it results that the approximations computed in Table 3 are always better or equal to the ones of Table 8. ∎

Differently from the preclusive rough approximation, if we consider the similarity rough approximation, we can see, comparing Tables 7 and 8 with Table 3, that the quality of the approximation is monotonic with respect to the quantity of information. This is a general result, as shown in the following proposition.

Proposition 2.7. *Let $\mathcal{K}^{(t_0)}(X)$ and $\mathcal{K}^{(t_1)}(X)$ be two incomplete information systems such that $\mathcal{K}^{(t_0)}(X) \preceq \mathcal{K}^{(t_1)}(X)$ (there is a monotonic increase of information). Then,*

$$(2.2) \qquad \forall H \subseteq X \quad \mathbf{L}_{\mathcal{R}}^{(t_0)}(H) \subseteq \mathbf{L}_{\mathcal{R}}^{(t_1)}(H) \subseteq H \subseteq \mathbf{U}_{\mathcal{R}}^{(t_1)}(H) \subseteq \mathbf{U}_{\mathcal{R}}^{(t_0)}(H).$$

Proof. Considering the similarity relation defined in equation (1), we have that if two objects are similar at time t_1 then they are also similar at time t_0: $x\mathcal{R}^{(t_1)}y$ implies $x\mathcal{R}^{(t_0)}y$. So the granule (similarity class) of an object x at time t_0 is bigger than the the granule (similarity class) of x at time t_1: $S^{(t_1)}(x) \subseteq S^{(t_0)}(x)$. Now, let H be a set of objects and compute its similarity rough approximation as defined in Equations (2a) and (2b). As regards to the lower approximation, we have that if $x \in L_{\mathcal{R}}^{(t_0)}(H)$, then $S^{(t_0)}(x) \subseteq H$. So, $S^{(t_1)}(x) \subseteq S^{(t_0)}(x) \subseteq H$, that is $x \in L_{\mathcal{R}}^{(t_1)}(H)$, and finally, $L_{\mathcal{R}}^{(t_0)}(H) \subseteq L_{\mathcal{R}}^{(t_1)}(H)$. Similarly, for the upper approximation, if $x \in U_{\mathcal{R}}^{(t_1)}(H)$, one obtains by definition that $S^{(t_1)}(x) \cap H \neq \emptyset$. Then $S^{(t_0)}(x) \cap H \neq \emptyset$, i.e., $x \in U_{\mathcal{R}}^{(t_0)}(H)$ and $U_{\mathcal{R}}^{(t_1)}(H) \subseteq U_{\mathcal{R}}^{(t_0)}(H)$. □

Concluding, if we suppose an increase of the information in time we have a monotononical behaviour of the similarity approximation and an unpredictable behaviour of the preclusive approximation, with the last one which is always better than the first one. From an intuitive point of view we can imagine a situation similar to the one drawn in Figure 2.
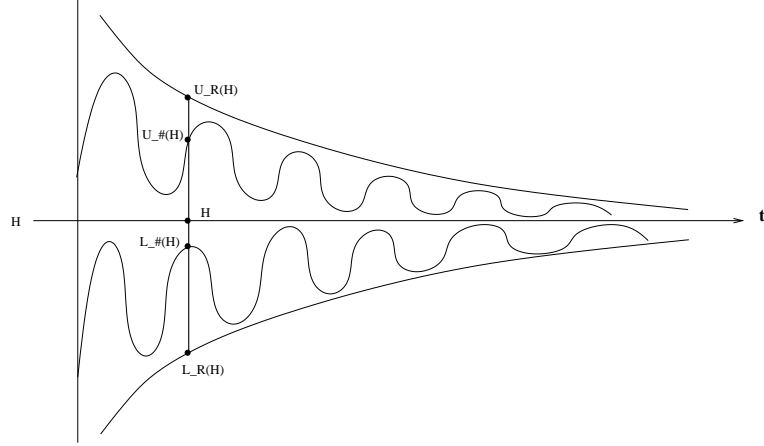
FIGURE 9. An imagininary representation of the time evolution of the similarity and preclusive rough approximations.

All the examples considered until now are about incomplete information systems and the similarity relation given by Equation (1). However, the pathological behavior remarked about the monotonic increase of information holds also in other contexts.

Another typical case of a reflexive and symmetric binary relation which in general is not transitive is the one induced by any distance. Precisely, if it can be introduced a metric assigning a distance between pairs of objects of the universe of an information system, we could say that two objects are similar if their distance is less than a fixed value ϵ. In this situation the similarity relation is sometimes called ϵ-*indiscernibility* and "the basic idea is that small differences on some attribute values may be judged insignificant" [SV95].

2.1. ϵ - **indiscernibility in presence of a metric.** Let $\mathcal{K}$ be an information system with numeric valued attributes, i.e., $Att(X) \subseteq \mathbb{R}$. Then, given an attribute $a \in Att(X)$ a way to define a pseudo–metric for a is

$$(2.3) \qquad d_a(x,y) := \frac{|F(x,a) - F(y,a)|}{\max\{F(z,a) : z \in X\} - \min\{F(w,a) : w \in X\}} \quad .$$

Clearly, this is a pseudo–metric and not a metric since it is not true that $d_a(x,y) = 0$ implies $x = y$ [Kel55]. Further, we remark that it is well defined since $\max\{F(z,a) : z \in X\} - \min\{F(w,a) : w \in X\}$ is never equal to 0 due to the assumption that for all attributes a there exist at least to objects x_1, x_2 such that $F(x_1,a) \neq F(x_2,a)$ (see Section 1).

In presence of two or more attributes collected in a set $\mathcal{D} \subseteq Att(X)$, we can introduce an overall distance:

$$d_D(x,y) := \sum_{a \in D} \omega_a d_a(x,y)$$

where ω_a are weights such that $\sum \omega_a = 1$ (a standard choice is $\omega_a = \frac{1}{|D|}$). So, for a fixed value $\epsilon \in [0,1]$ we can consider the similarity relation defined as follows:

$$(2.4) \qquad\qquad\qquad x\mathcal{R}_D^\epsilon y \quad \text{iff} \quad d_D(x,y) \leq \epsilon$$

The correlated preclusive relation is the following:

$$x \#_D^\epsilon y \quad \text{iff} \quad d_D(x, y) > \epsilon.$$

Let us stress that this is not the unique possibility to define a metric, for a more complete discussion see [NSS98].

Example 2.8. We apply now the similarity relation (12) on the same example used in [SV95] and [GMS98]. Let us consider the situation of 12 firms described in Table 9.

Firm	a_1	a_2	a_3	d
1	43	78	0	1
2	54	75	0	2
3	124	50	1	1
4	102	65	1	2
5	98	80	2	2
6	88	102	2	2
7	130	57	0	1
8	128	92	1	2
9	82	59	1	1
10	134	103	2	2
11	58	55	0	1
12	126	71	1	2

TABLE 9. Firms information table.

These firms are analyzed with respect to the following attributes:

- $a_1 =$ value of fixed capital,
- $a_2 =$ value of sales in the year preceding the application,
- $a_3 =$ kind of activity, classified according to the typology $0, 1$ and 2,

Further, these firms are divided in two classes: paid back their credit (d=1) or not (d=2).

In order to have a similarity relation we define a distance between two objects by considering only the first two attributes, i.e., $D = \{a_1, a_2\}$. So, by equation (12), the family of similarity relations, depending on ϵ, is:

$$x \mathcal{R}_D^\epsilon y \quad \text{iff} \quad \frac{1}{2} \left(\frac{|F(x, a_1) - F(y, a_1)|}{91} + \frac{|F(x, a_2) - F(y, a_2)|}{53} \right) \leq \epsilon$$

Let us set $\epsilon = 0.4$. We have, for example, that firm 1 is similar to firm 5:

$$d_D(1, 5) = \frac{1}{2} \left(\frac{|43 - 98|}{91} + \frac{|78 - 80|}{53} \right) \simeq 0.32 < 0.4 = \epsilon$$

On the other hand, firm 1 is preclusive to firm 12:

$$d_D(1, 12) = \frac{1}{2} \left(\frac{|43 - 126|}{91} + \frac{|78 - 71|}{53} \right) \simeq 0.52 > 0.4 = \epsilon$$

Further, firm 5 is similar to firm 12:

$$d_D(5, 12) = \frac{1}{2} \left(\frac{|98 - 126|}{91} + \frac{|71 - 80|}{53} \right) \simeq 0.24 < 0.4 = \epsilon$$

Thus, it is easily verified that $\mathcal{R}_D^{0,4}$ is not a transitive relation, since firm 1 is similar to firm 5 which is similar to firm 12, but firm 1 and firm 12 are not similar.

Now, let us consider the subsets of firms which paid back and did not pay back their credit: $P = \{1, 3, 7, 9, 11\}$ and $N = \{2, 4, 5, 6, 8, 10, 12\}$. In Table 10 we report their similarity and preclusive rough approximations, with respect to the relation $\mathcal{R}^{0.2}$.

$\mathbf{L}_{\mathcal{R}}(H)$	$\mathbf{L}_{\#}(H)$	H	$\mathbf{U}_{\#}(H)$	$\mathbf{U}_{\mathcal{R}}(H)$
{3,11}	{3,7,9,11}	P	{1,2,3,7,9,11,12}	{1,2,3,4,7,9,11,12}
{5,6,8,10}	{4,5,6,8,10}	N	{1,2,4,5,6,8,10,12}	{1,2,4,5,6,7,8,9,10,12}

TABLE 10. Example of rough approximations of the sets P and N.

As expected, preclusive approximations are better than the corresponding similarity ones. This is particularly evident when considering the boundary region, i.e., $\mathbf{U}(H)/\mathbf{L}(H)$, which represents the elements whose classification is uncertain. In fact, considering the set P we have that its preclusive boundary region $\mathbf{U}_{\#}(P)/\mathbf{L}_{\#}(P)$ is the set $\{1, 2, 12\}$, whereas its similarity boundary region $\mathbf{U}_{\mathcal{R}}(P)/\mathbf{L}_{\mathcal{R}}(P)$ is the set $\{1, 2, 4, 7, 9, 12\}$. Hence, the number of elements whose belonging to P is uncertain is half in the case of preclusive with respect to the corresponding similarity approximation.

Now, we turn our attention to open and closed sets. As can be seen in Table 11 the number of exact sets, both closed (and so also open) and clopen, decreases as the value of ϵ increases.

ϵ	Closed	Clopen
0.1	1024	1024
0.2	192	80
0.3	74	20
0.4	34	4
0.5	26	2
0.6	16	2
0.7	8	2
0.8	4	2
0.9	2	2

TABLE 11. Number of exact sets, depending on ϵ.

Thus, one could think that given $\epsilon_1 \leq \epsilon_2$, if H is closed with respect to $\#^{\epsilon_2}$, i.e., $H = \mathbf{U}_{\#}^{\epsilon_2}(H)$, then H is closed also with respect to $\#^{\epsilon_1}$, i.e., $H = \mathbf{U}_{\#}^{\epsilon_1}(H)$. In other words, one expects that $\epsilon_1 \leq \epsilon_2$ implies $\mathbb{C}^{\epsilon_2}(X) \subseteq \mathbb{C}^{\epsilon_1}(X)$. However, this is not true. For example, we have that $\{8, 10\} \in \mathbb{C}^{0.7}(X)$ but $\{8, 10\} \notin \mathbb{C}^{0.6}(X)$. The same can be said about clopen sets, in fact, we have that $\{5, 7, 8, 10\} \in \mathbb{CO}^{0.3}(X)$ but $\{5, 7, 8, 10\} \notin \mathbb{CO}^{0.2}(X)$.

When considering the rough approximation of a set H with respect to different relations relative to $\epsilon_1 \leq \epsilon_2$, we have the same unpredictability encountered in the case of incomplete information systems at different time steps $t_1 \leq t_2$. Indeed,

we can have that the approximation relative to ϵ_2 is better, worst or equal to the approximation relative to $\epsilon_1 \leq \epsilon_2$.

Let us refer to the information system of Table 9. We set $D = \{a_1, a_2\}$ and compute the approximation of some sets using the relation (12) with respect to two different values $\epsilon_1 = 0.6$ and $\epsilon_2 = 0.7$. The results are reported in Tables 12 and 13, respectively.

$\mathbf{L}_{\mathcal{R}}^{\epsilon_1}(H)$	$\mathbf{L}_{\#}^{\epsilon_1}(H)$	H	$\mathbf{U}_{\#}^{\epsilon_1}(H)$	$\mathbf{U}_{\mathcal{R}}^{\epsilon_1}(H)$
$\emptyset$	$\emptyset$	$\{8,10\}$	$\{6,8,10\}$	$\{1,2,3,4,5,6,7,8,9,10,12\}$
$\emptyset$	$\emptyset$	$\{2\}$	$\{1,2\}$	$\{1,2,4,5,6,7,8,9,11,12\}$
$\emptyset$	$\emptyset$	$\{2,4,5\}$	X	X

TABLE 12. Example of approximation for $\epsilon_1 = 0.6$

$\mathbf{L}_{\mathcal{R}}^{\epsilon_2}(H)$	$\mathbf{L}_{\#}^{\epsilon_2}(H)$	H	$\mathbf{U}_{\#}^{\epsilon_2}(H)$	$\mathbf{U}_{\mathcal{R}}^{\epsilon_2}(H)$
$\emptyset$	$\emptyset$	$\{8,10\}$	$\{8,10\}$	$\{1,2,3,4,5,6,7,8,9,10,12\}$
$\emptyset$	$\emptyset$	$\{2\}$	$\{1,2,9,11\}$	$\{1,2,3,4,5,6,7,8,9,11,12\}$
$\emptyset$	$\emptyset$	$\{2,4,5\}$	X	X

TABLE 13. Example of approximation for $\epsilon_2 = 0.7$

Thus, in the considered three cases we have all the possible situations: the preclusive approximation relative to ϵ_1 is better (case $\{2\}$), worst (case $\{8, 10\}$) or equal (case $\{2, 4, 5\}$) of the approximation relative to ϵ_2. ∎

A particular situation that guarantees us to have a better approximation going from ϵ_1 to ϵ_2 is given by a monotonic (with respect to ϵ) increase of knowledge.

Proposition 2.9. *Let $\mathcal{K}(X)$ be a real valued information system (i.e., $val(X) \subseteq \mathbb{R}$) and let $\epsilon_1, \epsilon_2 \in [0, 1]$. If, with respect to the similarity relation (12), it happens that $\mathbb{C}^{\epsilon_2}(X) \subseteq \mathbb{C}^{\epsilon_1}(X)$ then*

$$\forall H \subseteq X \quad \mathbf{L}_{\#}^{\epsilon_2}(H) \subseteq \mathbf{L}_{\#}^{\epsilon_1}(H) \subseteq H \subseteq \mathbf{U}_{\#}^{\epsilon_1}(H) \subseteq \mathbf{U}_{\#}^{\epsilon_2}(H).$$

Proof. Same as Proposition 2.5. □

Example 2.10. Let us refer to the Example 2.8. If we choose $\epsilon_1 = 0.7$ and $\epsilon_2 = 0.8$ we have that $\mathbb{C}^{\epsilon_2}(X) \subseteq \mathbb{C}^{\epsilon_1}(X)$, as can be seen by Hasse diagrams of figure 10 and 11.

So we expect that the approximations relative to $\epsilon_1 = 0.7$ are better than the ones relative to $\epsilon_2 = 0.8$. As an example we compute the approximations of the same sets of Example 2.8 in the case of $\epsilon_2 = 0.8$.

By comparing Table 13 with Table 14, it is easily seen that Proposition 2.9 is verified. ∎

On the other hand, in the case of the similarity rough approximation, we always have a monotonic behavior.
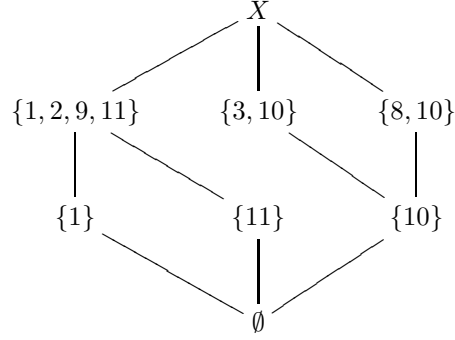
FIGURE 10. Closed sets: case $\epsilon_1 = 0.7$.
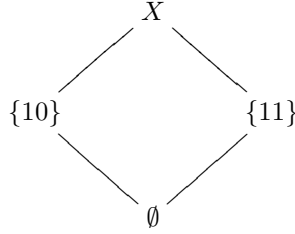
FIGURE 11. Closed sets: case $\epsilon_2 = 0.8$.

$\mathbf{L}_{\mathcal{R}}^{\epsilon}(H)$	$\mathbf{L}_{\#}^{\epsilon}(H)$	H	$\mathbf{U}_{\#}^{\epsilon}(H)$	$\mathbf{U}_{\mathcal{R}}^{\epsilon}(H)$
$\emptyset$	$\emptyset$	{8,10}	X	X
$\emptyset$	$\emptyset$	{2}	X	X
$\emptyset$	$\emptyset$	{2,4,5}	X	X

TABLE 14. Example of approximation, $\epsilon = 0.8$

Proposition 2.11. *Let $\mathcal{K}(X)$ be a real valued information system (i.e., $val(X) \subseteq \mathbb{R}$) and let $\epsilon_1, \epsilon_2 \in [0,1]$ with $\epsilon_1 \leq \epsilon_2$. Then, once chosen a set of attributes $D \subseteq Att(X)$ and defined the similarity relation $\mathcal{R}_D^{\epsilon}$ as in equation (12), the following holds*

$$\mathbf{L}_{\mathcal{R}}^{\epsilon_2}(H) \subseteq \mathbf{L}_{\mathcal{R}}^{\epsilon_1}(H) \subseteq H \subseteq \mathbf{U}_{\mathcal{R}}^{\epsilon_1}(H) \subseteq \mathbf{U}_{\mathcal{R}}^{\epsilon_2}(H)$$

Proof. The proof is similar to the one of proposition 2.7, considering that if $\epsilon_1 \leq \epsilon_2$ than $x\mathcal{R}^{\epsilon_1}y$ implies $x\mathcal{R}^{\epsilon_2}y$. $\qquad\square$

3. Quasi BZ Distributive Lattices

In the previous sections, through concrete cases based on the power set of a universe, we introduced preclusivity spaces and the rough approximations generated by them. Moreover, it emerged that in all the considered examples the preclusive

approximation was always better than the corresponding similarity approximation. The aim of this section is to give an algebraic characterization to preclusivity spaces in terms of quasi Brouwer Zadeh distributive lattice structures ([Cat97, Cat98]). This will also lead to a theoretical proof of the better behavior of the preclusive rough approximation with respect to the similarity one.

In particular, we considered the concrete algebra $\langle \mathcal{P}(X), \cup, \cap, {}^c, {}^{\#} \rangle$ based on the power set $\mathcal{P}(X)$ of a universe X endowed with the set theoretic union, intersection, complement operations, and a unusual complementation $\#$.

Now, through an abstraction, we are going to introduce an abstract system consisting of a set Σ, corresponding to the power set $\mathcal{P}(X)$ of the concrete example, upon which the following operations are defined:

- two binary operations, $\vee$ and $\wedge$, corresponding respectively to the set theoretic union and intersection of the concrete example;
- two unary operations, $'$ and $\sim$, corresponding to the standard and preclusive complementations of the concrete example.

Definition 3.1. A system $\langle \Sigma, \wedge, \vee, {}', {}^{\sim}, 0, 1 \rangle$ is a *quasi Brouwer Zadeh (BZ) distributive lattice* if the following properties hold:

(1) Σ is a distributive lattice with respect to the join and the meet operations $\vee, \wedge$ whose induced partial order relation is

$$a \leq b \quad \text{iff} \quad a = a \wedge b \quad (\text{equivalently } b = a \vee b)$$

Moreover, it is required that Σ is bounded by the least element 0 and the greatest element 1:

$$\forall a \in \Sigma, \ 0 \leq a \leq 1$$

(2) The unary operation $' : \Sigma \mapsto \Sigma$ is a *Kleene* (also *Zadeh* or *fuzzy*) complementation. In other words for arbitrary $a, b \in \Sigma$:
 - (K1) $\quad a'' = a$
 - (K2) $\quad (a \vee b)' = a' \wedge b'$
 - (K3) $\quad a \wedge a' \leq b \vee b'$.

(3) The unary operation $\sim : \Sigma \mapsto \Sigma$ is a *Brouwer* (or *intuitionistic*) complementation. In other words for arbitrary $a, b \in \Sigma$:
 - (B1) $\quad a \wedge a^{\sim\sim} = a$
 - (B2) $\quad (a \vee b)^{\sim} = a^{\sim} \wedge b^{\sim}$
 - (B3) $\quad a \wedge a^{\sim} = 0$

(4) The two complementations are linked by the interconnection rule which must hold for arbitrary $a \in \Sigma$:
 - (in) $\quad a^{\sim} \leq a'$

We remark that under condition (K1), property (K2) is equivalent to the following

(K2a) $\quad (a \wedge b)' = a' \vee b'$;
(K2b) $\quad a \leq b$ implies $b' \leq a'$;
(K2c) $\quad b' \leq a'$ implies $a \leq b$.

In general neither the non contradiction law $a \wedge a' = 0$ nor the excluded middle law $a \vee a' = 1$ for the Kleene negation are required to hold.

As to the intuitionistic negation, we have that under condition (B1), condition (B2) is equivalent only to the weak form of non contraposition law

(B2a) $\quad a \leq b$ implies $b^{\sim} \leq a^{\sim}$.

In general the strong contraposition law, $b^\sim \leq a^\sim$ implies $a \leq b$, and the dual de Morgan law, $(a \wedge b)^\sim = a^\sim \vee b^\sim$ do not hold for the Brouwer negation. As expressed by property (B3), the non contradiction law is satisfied by all elements, whereas, in general, the excluded middle law is not required to hold.

In the framework of quasi BZ lattices, one can naturally introduce the *anti-Brouwer complement* $^\flat : \Sigma \mapsto \Sigma$ defined for every $a \in \Sigma$ as: $a^\flat := a'^{\sim'}$. It can be easily shown that the operation $^\flat$ satisfies the following properties:

(AB1) $a \leq b$ implies $b^\flat \leq a^\flat$ [Contraposition law]

(AB2) $a \vee a^\flat = 1$ [Excluded middle law]

(AB3) $a^{\flat\flat} \leq a$.   [Weak double negation law]

Our three complements turn out to be connected by the following chain:

$$\forall a \in \Sigma , \quad a^\sim \leq a' \leq a^\flat .$$

Proposition 3.2. *Let $\langle \Sigma, \wedge, \vee, ', \sim, 0, 1 \rangle$ be a quasi BZ distributive lattice. Then, the map $\mathbf{i} : \Sigma \to \Sigma$ such that $\mathbf{i}(a) := a^{\flat\flat} = a'^{\sim\sim'}$ is an interior operator. That is, the following are satisfied:*

(I_0)	$1 = \mathbf{i}(1)$	(*normalized*)
(I_1)	$\mathbf{i}(a) \leq a$	(*decreasing*)
(I_2)	$\mathbf{i}(a) = \mathbf{i}(\mathbf{i}(a))$	(*idempotent*)
(I_3)	$\mathbf{i}(a \wedge b) \leq \mathbf{i}(a) \wedge \mathbf{i}(b)$	(*submultiplicative*)

Dually, the map $\mathbf{c} : \Sigma \to \Sigma$ such that $\mathbf{c}(a) := a^{\sim\sim}$ is a closure operator, i.e.:

(C_0)	$0 = \mathbf{c}(0)$	(*normalized*)
(C_1)	$a \leq \mathbf{c}(a)$	(*increasing*)
(C_2)	$\mathbf{c}(a) = \mathbf{c}(\mathbf{c}(a))$	(*idempotent*)
(C_3)	$\mathbf{c}(a) \vee \mathbf{c}(b) \leq \mathbf{c}(a \vee b)$	(*subadditive*)

Under condition (I_0)–(I_2) the submultiplicative property (I_3) is equivalent to the monotonicity condition $a \leq b$ implies $\mathbf{i}(a) \leq \mathbf{i}(b)$. Dually, under condition (C_0)–(C_2) the subadditive property (C_3) is equivalent to the monotonicity condition $a \leq b$ implies $\mathbf{c}(a) \leq \mathbf{c}(b)$.

Since in general $\mathbf{i}(a) \leq a$, an element $e \in \Sigma$ is said to be *innerdefinable (open)* iff $\mathbf{i}(e) = e$. Analogously, since in general $a \leq \mathbf{c}(a)$, an element $f \in \Sigma$ is said to be *outerdefinable (closed)* iff $f = \mathbf{c}(f)$. We will indicate by $\mathbb{O}(\Sigma)$ the set of all *innerdefinable (open)* elements of the space, while $\mathbb{C}(\Sigma)$ will represent the set of all *outerdefinable (closed)* elements. In other words:

$$\mathbb{O}(\Sigma) = \{e \in \Sigma : e = \mathbf{i}(e)\}$$
$$\mathbb{C}(\Sigma) = \{f \in \Sigma : f = \mathbf{c}(f)\}$$

It is trivial to prove that an element a is closed iff a' is open, and vice versa an element a is open iff a' is closed. In this way there is a duality between open and closed elements determined by the Kleene negation:

$$\mathbb{O}(\Sigma) = \{e \in \Sigma : e' \in \mathbb{C}(\Sigma)\}$$
$$\mathbb{C}(\Sigma) = \{f \in \Sigma : f' \in \mathbb{O}(\Sigma)\}$$

In general, these two subsets of Σ do not coincide, neither one is a subset of the other. Thus, it is worthwhile to consider also the set of all *innouter* (or *clopen*) elements:

$$\mathbb{CO}(\Sigma) := \mathbb{O}(\Sigma) \cap \mathbb{C}(\Sigma).$$

Moreover, $\mathbb{CO}(\Sigma)$ is not empty because it contains the least element 0 and the greatest element 1 of the lattice Σ. In the sequel we are also interested to the so–called *exterior* of an element $a \in \Sigma$ defined as the open element $\mathbf{e}(a) := \mathbf{c}(a)'$.

The above considerations lead to the definition of an abstract approximation space generated by a quasi BZ lattice.

Definition 3.3. Let $\langle \Sigma, \leq, ', {}^{\sim}, 0, 1 \rangle$ be a quasi BZ distributive lattice. The induced rough approximation space is the structure

$$\langle \Sigma, \mathbb{O}(\Sigma), \mathbb{C}(\Sigma), \mathbf{i}, \mathbf{c} \rangle$$

where

- Σ is the set of approximable elements;
- $\mathbb{O}(\Sigma) \subseteq \Sigma$ is the set of *innerdefinable* elements, such that 0 and $1 \in \mathbb{O}(\Sigma)$;
- $\mathbb{C}(\Sigma) \subseteq \Sigma$ is the set of *outerdefinable* elements, such that 0 and $1 \in \mathbb{C}(\Sigma)$;
- $\mathbf{i} : \Sigma \to \mathbb{O}(\Sigma)$ is the *inner approximation map*;
- $\mathbf{c} : \Sigma \to \mathbb{C}(\Sigma)$ is the *outer approximation map*;

For any element $a \in \Sigma$, its *rough approximation* is defined as the pair:

$$r(a) := \langle \mathbf{i}(a), \mathbf{c}(a) \rangle \quad [\text{with} \quad \mathbf{i}(a) \leq a \leq \mathbf{c}(a)]$$

drawn in the following diagram:

This approximation is the best approximation by open and closed elements that it is possible to define on a quasi BZ structure. To be precise, for any element $a \in X$ the following holds:

- (L1) $\mathbf{i}(a)$ is an open element ($\mathbf{i}(a) \in \mathbb{O}(X)$);
- (L2) $\mathbf{i}(a)$ is an open *lower* approximation of a ($\mathbf{i}(a) \leq a$);
- (L3) $\mathbf{i}(a)$ is the best lower approximation of a by open elements (let $e \in \mathbb{O}(X)$ be such that $e \leq a$, then $e \leq \mathbf{i}(a)$).

By properties (L1)–(L3), it follows that the interior of an element a can be expressed as

$$\mathbf{i}(a) = \max\{x \in \mathbb{O}(X) : x \leq a\}$$

Analogously, for the closure operator,

- (U1) $\mathbf{c}(a)$ is a closed element ($\mathbf{c}(a) \in \mathbb{C}(X)$).
- (U2) $\mathbf{c}(a)$ is a closed *upper* approximation of a ($a \leq \mathbf{c}(a)$).
- (U3) $\mathbf{c}(a)$ is the best outer approximation of a by closed elements (let $f \in \mathbb{C}(X)$ be such that $a \leq f$, then $\mathbf{c}(a) \leq f$).
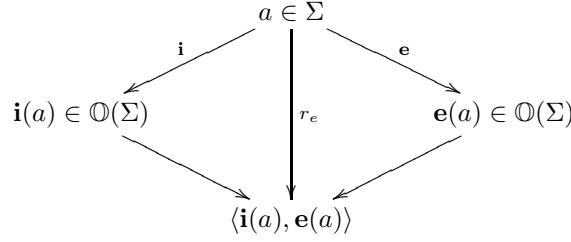
By properties (U1)–(U3), it follows that the closure of an element a can be expressed as

$$\mathbf{c}(a) = \min\{y \in \mathbb{C}(X) : a \leq y\}$$

An equivalent way to define a rough approximation is to consider instead of the interior–closure pair the interior–exterior pair:

$$r_e(a) := \langle \mathbf{i}(a), \mathbf{e}(a) \rangle := \langle \mathbf{i}(a), \mathbf{c}(a)' \rangle$$

drawn in the following diagram

$$a \in \Sigma$$

$$\mathbf{i}(a) \in \mathbb{O}(\Sigma) \qquad r_e \qquad \mathbf{e}(a) \in \mathbb{O}(\Sigma)$$

$$\langle \mathbf{i}(a), \mathbf{e}(a) \rangle$$

We remark that both $\mathbf{i}$(a) and $\mathbf{e}$(a) are open elements and that given the interior – exterior approximation $r_e(a)$, the interior – closure approximation $r(a)$ can be obtained in an obvious way, and vice versa.

Further, the interior and closure operators satisfy some typical modal properties, once interpreted the interior as a necessity operator and the closure as a possibility operator ([Che88]).

Proposition 3.4. *In any quasi BZ distributive lattice the following conditions hold:*

(mod–1p)

$$\mathbf{i}(1) = 1.$$

That is: if a sentence is true, then also its necessity its true (necessitation rule).

(mod–2p)

$$\mathbf{i}(a) \leq a \leq \mathbf{c}(a).$$

In other words: necessity implies actuality and actuality implies possibility (a characteristic principle of the modal system T).

(mod–3p)

$$\mathbf{i}(a) = \mathbf{i}(\mathbf{i}(a)); \quad \mathbf{c}(a) = \mathbf{c}(\mathbf{c}(a)).$$

necessity and possibility are idempotent operators (a characteristic S_4–principle).

As a consequence of these properties we can say that the interior and closure operators induced by a quasi BZ distributive lattice, present a typical behaviour of a **T** system, based on a Kleene lattice, instead of on a Boolean algebra, satisfying also the S_4–principle.

Generally, the following properties of modalities are not satisfied, as showed in Example 3.5.

(B) $a \leq \mathbf{i}(\mathbf{c}(a)$ (a characteristic B-principle);
(S_5) $\mathbf{i}(a) = \mathbf{c}(\mathbf{i}(a))$, $\mathbf{c}(a) = \mathbf{i}(\mathbf{c}(a))$ (a characteristic S_5-principle).

$$1 = 0' = 0^\sim$$

$$b' = c \qquad\qquad d = a'$$

$$d' = b^\sim = a \qquad\qquad b = a^\sim = c'$$

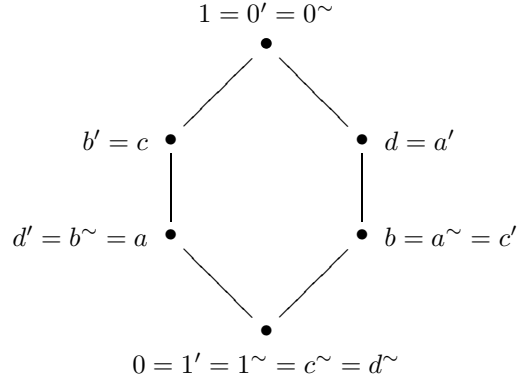$$0 = 1' = 1^\sim = c^\sim = d^\sim$$

FIGURE 12

Example 3.5. Let us consider the quasi BZ distributive lattice given in the Hasse diagram of Figure 12.

Then, we have that b is a closed element, i.e., $b = \mathbf{c}(b)$ and $\mathbf{i}(\mathbf{c}(b)) = \mathbf{i}(b) = 0$. So, $\mathbf{c}(b) \neq \mathbf{i}(\mathbf{c}(b))$ and $\mathbf{i}(\mathbf{c}(b)) < b$. Further, c is an open element, i.e., $\mathbf{i}(c) = c$ and $\mathbf{c}(\mathbf{i}(c)) = 1$. So $\mathbf{i}(c) \neq \mathbf{c}(\mathbf{i}(c))$. ∎

Even if the interior and closure operators give the best approximation by closed and open elements, they are not the only possible ones. Another interesting possibility is given by the following two operators:

(3.1a) $$\nu : \Sigma \to \mathbb{C}(\Sigma) \quad \nu(a) := a'^\sim$$

(3.1b) $$\mu : \Sigma \to \mathbb{O}(\Sigma) \quad \mu(a) := a^{\sim\prime}$$

These mappings can be considered approximation operators because the following order chain $\nu(a) \leq a \leq \mu(a)$ holds. Moreover, they behave as *modal operators* ([Che88]), as shown in the following proposition.

Proposition 3.6. *In any quasi BZ distributive lattice the following conditions hold:*

(mod–1s)
$$\nu(1) = 1.$$

That is: if a sentence is true, then also its necessity its true (necessitation rule).

(mod–2s)
$$\nu(a) \leq a \leq \mu(a).$$

In other words: necessity implies actuality and actuality implies possibility (a characteristic principle of the modal system T).

(mod–3s)
$$a \leq \nu(\mu(a)).$$

Actuality implies necessity of possibility (a characteristic B-principle).

As a consequence of these properties we can say that the modal operators induced by a quasi BZ distributive lattice, present a typical behavior of a **B** system, based on a Kleene lattice, instead of on a Boolean algebra.

Generally, ν and μ do not satisfy the S_4 and S_5 principles:

(S_4) $\nu(a) = \nu(\nu(a)); \quad \mu(a) = \mu(\mu(a))$ (a characteristic S_4-principle).

(S_5) $\nu(a) = \mu(\nu(a))$, $\mu(a) = \nu(\mu(a))$ (a characteristic S_5-principle).

Example 3.7. Let us consider the same quasi BZ lattice of Example 3.5. Then, $\nu(c) = a$, so $\nu(c) \neq \nu(\nu(c)) = 0$ and $\nu(c) \neq \mu(\nu(c)) = c$. Further, $\mu(a) = c$, so $\mu(a) \neq \mu(\mu(a)) = 1$ and $\mu(a) \neq \nu(\mu(a)) = a$. ∎

However, necessity and possibility are not outer and inner operators. In fact, in general, they do not satisfy the idempotent properties, but only the following weaker form:

$$\nu(\nu(a)) \leq \nu(a)$$
$$\mu(a) \leq \mu(\mu(a))$$

Example 3.8. Let us consider the same quasi BZ lattice of Example 3.5. Then, $\mu(a) = c < 1 = \mu(\mu(a))$ and $\nu(\nu(d)) = 0 < b = \nu(d)$. ∎

The major drawback of this approximation is that the rough approximation based on modal operators is worst than the corresponding one based on the interior - closure operators, in the sense that it captures less information about the approximable elements. In fact, for any element $a \in X$, the following order chain holds:

(3.2) $$\nu(a) \leq \mathbf{i}(a) \leq a \leq \mathbf{c}(a) \leq \mu(a).$$

3.1. **Rough Approximation Spaces as quasi BZ Lattices.** We have seen that given an element a of a quasi BZ distributive lattice, a possible rough approximation of a is given by $r_e(a) = \langle \mathbf{i}(a), \mathbf{e}(a) \rangle$. Now, we show that it is possible to give to the collection of all rough approximations, $r_e(a)$, the structure of quasi BZ distributive lattice.

Let $\langle \Sigma, \wedge, \vee, ', {}^{\sim}, 0, 1 \rangle$ be a quasi BZ lattice. For the sake of simplicity, we define

$$\forall a \in \Sigma \quad a_i := \mathbf{i}(a) \quad a_e := \mathbf{e}(a) = \mathbf{c}(a)'$$

We recall that both a_i and a_e are open elements: $a_i, a_e \in \mathbb{O}(\Sigma)$. Further, we define the collection of all rough approximations as the set $\mathbb{A} := \{\langle a_i, a_e \rangle : a \in \Sigma\}$. Then, the following proposition holds.

Proposition 3.9. *The structure* $\langle \mathbb{A}, \sqcap, \sqcup, {}^-, {}^{\approx}, \langle 0, 1 \rangle, \langle 1, 0 \rangle \rangle$ *where:*

$$\langle a_i, a_e \rangle \sqcap \langle b_i, b_e \rangle := \langle \mathbf{i}(a_i \wedge b_i), a_e \vee b_e \rangle$$
$$\langle a_i, a_e \rangle \sqcup \langle b_i, b_e \rangle := \langle a_i \vee b_i, \mathbf{i}(a_e \wedge b_e) \rangle$$
$$\langle a_i, a_e \rangle^- := \langle a_e, a_i \rangle$$
$$\langle a_i, a_e \rangle^{\approx} := \langle \mathbf{e}((a_e)^{\flat}), (a_e)^{\flat} \rangle$$

is a quasi BZ distributive lattice.

The partial order induced by the lattice operators, $\sqcap$ and $\sqcup$, is

$$\langle a_i, a_e \rangle \sqsubseteq \langle b_i, b_e \rangle \quad \text{iff} \quad a_i \leq b_i \text{ and } b_e \leq a_e$$

In this environment the interior, closure and exterior operators are respectively defined as:

$$\mathbf{I}(\langle a_i, a_e \rangle) := \langle \mathbf{i}(a_i), \mathbf{e}(\mathbf{i}(a_i)) \rangle = \langle a_i, \mathbf{e}(a_i) \rangle$$
$$\mathbf{C}(\langle a_i, a_e \rangle) := \langle \mathbf{e}(\mathbf{i}(a_e)), \mathbf{i}(a_e) \rangle = \langle \mathbf{e}(a_e), a_e \rangle$$
$$\mathbf{E}(\langle a_i, a_e \rangle) := \langle a_e, \mathbf{e}(a_e) \rangle$$

The collection of open and closed sets are respectively

$$\mathbb{O}(\mathbb{A}) = \{\langle a_i, \mathbf{e}(a_i)\rangle : a \in \Sigma\}$$
$$\mathbb{C}(\mathbb{A}) = \{\langle \mathbf{e}(b_e), b_e\rangle : b \in \Sigma\}$$

3.2. Preclusivity Spaces as quasi BZ distributive lattices.

Now, we come back to preclusivity spaces, and analize them in a quasi BZ lattice perspective.

Let $\mathcal{S} = \langle X, \# \rangle$ be a preclusivity space as defined in Definition 1.9. As recalled in Proposition 1.10, the structure $\langle \mathcal{P}(X), \cap, \cup, {}^c, \#, \emptyset, X \rangle$, is a quasi BZ distributive lattice. Thus, the interior $\mathbf{L}_\#(H) = H^{c\#\#c}$ and closure $\mathbf{U}_\#(H) = H^{\#\#}$ operators as defined in Proposition 3.2 are just the preclusive lower and upper approximation as defined in Propositions 1.12 and 1.14. More interestingly, when considering the modal operators ν and μ of equations (13) it happens that they correspond to the similarity approximations given in Definition 1.7 by equation (2): $\nu(H) = H^{c\#} = \mathbf{L}_\mathcal{R}(H)$ and $\mu(H) = H^{\#c} = \mathbf{U}_\mathcal{R}(H)$.

As an application of the order chain (14) to the present case we have that $\forall H \subseteq X$:

$$\mathbf{L}_\mathcal{R}(H) \subseteq \mathbf{L}_\#(H) \subseteq H \subseteq \mathbf{U}_\#(H) \subseteq \mathbf{U}_\mathcal{R}(H)$$

This means that the preclusive rough approximation of a set H is always closer to H than the one given by similarity. Thus, what was just a conjecture in Example 1.16 has now been proved.

Now, we can apply Proposition 3.9 and give the structure of quasi BZ lattices to the collection of all preclusive rough approximations. Given a set of objects H the pair interior–exterior $\langle \mathbf{i}(H), \mathbf{e}(H)\rangle$ is identified with the two sets: $\langle H_i, H_e\rangle := \langle \mathbf{L}_\#(H), (\mathbf{U}_\#(H))^c\rangle = \langle H^{c\#\#c}, H^{\#\#c}\rangle$. So, we have that the collection of all such pairs $\{\langle H_i, H_e\rangle : H \subseteq \mathcal{P}(X)\}$ is a quasi BZ distributive lattice, once defined the operators:

$$\langle H_i, H_e\rangle \sqcap \langle K_i, K_e\rangle := \langle \mathbf{L}_\#(H_i \cap K_i), H_e \cup K_e\rangle$$
$$\langle H_i, H_e\rangle \sqcup \langle K_i, K_e\rangle := \langle H_i \cup K_i, \mathbf{L}_\#(H_e \cap K_e)\rangle$$
$$\langle H_i, H_e\rangle^- := \langle H_e, H_i\rangle$$
$$\langle H_i, H_e\rangle^\approx := \langle (\mathbf{U}_\#((H_e)^\flat))^c, (H_e)^\flat\rangle$$

4. BZ Lattices and Pawlak Rough Sets

So far we have considered concrete quasi-BZ lattices and concrete rough approximation spaces induced by a preclusivity space $\langle X, \#\rangle$, where the preclusive relation $\#$ is irreflexive and symmetric (or, equivalently, by a similarity space $\langle X, \mathcal{R}\rangle$, where the similarity relation is reflexive and symmetric). Now, in order to obtain Pawlak's rough approximation spaces ([Paw81, Paw82, Paw92], it will be sufficient to assume the stronger assumption according to which in any similarity space $\langle X, \mathcal{R}\rangle$ the relation $\mathcal{R}$ is an equivalence relation (reflexive, symmetric and transitive). As a consequence, $\mathcal{R}$ will determine a partition of the universe X into a set of equivalence classes E_i: two elements x and y will be undistinguishable ($x\mathcal{R}y$) iff they belong to one and the same equivalence class E_i. Given an element $x \in X$ we can define the *equivalence class* generated by x as the set

$$E(x) = \{y \in X : x\mathcal{R}y\}$$

Obviously, $E(x)$ is not empty because x belongs to it and, analogously to the similarity case, it constitute a *granule* of knowledge about x. Let us stress that two granule are either disjoint or equal between them.

In particular, in the classical Pawlak's approach to rough sets the equivalence relation is based on the equality of attribute values of a *complete* information system $\mathcal{K}(X)$. Precisely, fixed a subset of attributes $D \subseteq Att(X)$ the equivalence relation of any two objects is defined as:

(4.1) $x \equiv y \quad \text{iff} \quad \forall a \in D\ F(x,a) = F(y,a)$

Obviously, the corresponding preclusive relation is defined as:

(4.2) $x \not\equiv y \quad \text{iff} \quad \exists a \in D\ F(x,a) \neq F(y,a)$

Following the general theory of Section 3, we have that the structure $\mathbb{P} = \langle \mathcal{P}(X), \cap, \cup, ^c, \neq, \emptyset, X \rangle$ is a quasi BZ distributive lattice. However, the transitive property of the similarity relation introduces some strong "classical" feature.

Example 4.1. Let us consider the flat information system defined in Example 1.2 and reported below.

Flat	Price	Rooms	Down-Town	Furniture
f_1	high	2	yes	no
f_2	high	1	yes	no
f_3	high	2	yes	no
f_4	low	1	no	no
f_5	low	1	no	no
f_6	medium	1	yes	yes

We set $D = Att(X)$ and $H = \{f_3, f_4\}$. Computing the similarity rough approximation $\langle \nu(H), \mu(H) \rangle$ and the preclusivity rough approximation $\langle \mathbf{i}(H), \mathbf{c}(H) \rangle$, we have that

$$\langle \nu(H), \mu(H) \rangle = \langle \emptyset, \{f_1, f_3, f_4, f_5\} \rangle = \langle \mathbf{i}(H), \mathbf{c}(H) \rangle$$

that is, the two approximations coincide. So, in particular, we have that

$$\nu(H) = \mathbf{i}(H)$$

that is, the necessity of H is equivalent to the interior of H, and

$$\mu(H) = \mathbf{c}(H)$$

that is, the possibility of H is equivalent to the closure of H. Further, the following properties hold: $\nu(\nu(H)) = \emptyset = \nu(H)$ and $\mu(\mu(H)) = \{f_1, f_3, f_4, f_5\} = \mu(H)$, that is, the modal operators ν and μ are idempotent. ∎

The properties analyzed in the previous example, i.e, idempotency of necessity and possibility, equivalence of necessity (resp., possibility) and interior (resp., closure) operators, are true for any element $H \in \mathcal{P}(X)$. This is due to the fact that the system $\mathbb{P}$ satisfies the axioms of a stronger structure than quasi BZ lattices: BZ lattices.

Definition 4.2. A structure $\langle \Sigma, \wedge, \vee, ', ^\sim, 0, 1 \rangle$ is a *BZ distributive lattice* if it is a quasi BZ distributive lattice satisfying the stronger interconnection rule:

 (s–in) $a^{\sim\sim} = a^{\sim\prime}$

Definition 4.3. A BZ distributive lattice satisfying also the $\vee$ de Morgan property

(B2a) $(a \wedge b)^\sim = a^\sim \vee b^\sim$

is called a *de Morgan BZ (BZ^{dM}) distributive lattice.*

A consequence of the new axiom (s–in) is that the necessity mapping coincide with the interior operator and the possibility mapping with the closure operator.

Proposition 4.4. *Let $\langle \Sigma, \wedge, \vee, ', ^\sim, 0, 1 \rangle$ be a BZ distributive lattice. Then, the following are true:*

$$\forall a \in \Sigma \quad \nu(a) = a'^\sim = a'^{\sim\sim'} = \mathbf{i}(a)$$
$$\forall a \in \Sigma \quad \mu(a) = a^{\sim'} = a^{\sim\sim} = \mathbf{c}(a)$$

Further, we have that the modal operators, ν and μ, have the stronger behaviour of an S_5–modal like system, always based on a Kleene algebra instead of on a Boolean one. That is, beside properties (mod–1)–(mod–3), the following are satisfied:

(mod–4)

$$\nu(\nu(a)) = \nu(a)$$
$$\mu(\mu(a)) = \mu(a)$$

Necessity of necessity is equal to necessity; similarly for possibility (a characteristic S_4-principle).

(mod–5)

$$\mu(a) = \nu(\mu(a))$$
$$\nu(a) = \mu(\nu(a))$$

Possibility is equal to the necessity of possibility; whereas necessity is equal to the possibility of necessity (a characteristic S_5-principle).

Finally, another consequence of axiom (s–in) is the equivalence between the collection of open and closed sets:

$$\mathbb{O}(\Sigma) = \{a \in \Sigma : a = \mathbf{i}(a)\} = \{a \in \Sigma : a = \mathbf{c}(a)\} = \mathbb{C}(\Sigma)$$

So, we have not to distinguish between upper exact and lower exact sets, we only have exact sets, i.e., sets which are clopen. Let us indicate the collection of exact sets as Σ_e.

We remark that this is not the only way to define exact elements. We have seen that in general the order chain $\nu(a) \leq a \leq \mu(a)$ does not hold. Clearly, this is a fuzzy situation, since in a classical environment we would have no difference among necessity, actuality and possibility, i.e., we are interested to those elements for which $e = \mu(e)$ (equivalently, $e = \nu(e)$). This leads us to define the substructure of all *Modal sharp (M-sharp)* (exact, crisp) elements, denoted by $\Sigma_{e,M}$, as follows:

$$\Sigma_{e,M} := \{e \in \Sigma : \mu(e) = e\} = \{e \in \Sigma : \nu(e) = e\}$$

However, due to the strong interconnection holding between the two orthocomplementations in a BZ structure, the collection of M–sharp elements coincide with Σ_e. Further, since in general $a \wedge a' \neq 0$ (equivalently, $a \vee a' \neq 1$) it is possible to consider as *Kleene sharp (K-sharp)* the elements which satisfy the non contradiction (or equivalently the excluded middle) law with respect to the Kleene negation:

$$\Sigma_{e,'} := \{e \in \Sigma : e \wedge e' = 0\} = \{e \in \Sigma : e \vee e' = 1\}$$

In a BZ^{dM} distributive lattice also this set coincide with Σ_e, otherwise, in a BZ distributive lattice it holds the relation $\Sigma_e \subseteq \Sigma_{e,'}$.

Summarizing, we have that in a BZ distributive lattice, we can give only one rough approximation through exact sets. The approximation operators turn out to be respectively an interior and a closure operator with the further property that they have a typical S_5 modal behaviour.

Definition 4.5. Let $\langle \Sigma, \leq, ', \sim, 0, 1 \rangle$ be a BZ distributive lattice. The induced rough approximation space is the structure
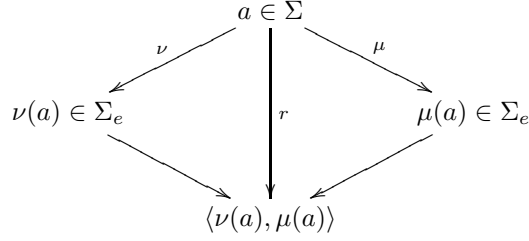
$$\langle \Sigma, \Sigma_e, \nu, \mu \rangle$$

where

- Σ is the set of approximable elements;
- $\Sigma_e \subseteq \Sigma$ is the set of exact elements;
- $\nu : \Sigma \to \Sigma_e$ is the *inner approximation map*;
- $\mu : \Sigma \to \Sigma_e$ is the *outer approximation map*;

For any element $a \in \Sigma$, its *rough approximation* is defined as the pair:

$$r(a) := \langle \nu(a), \mu(a) \rangle \quad [\text{with} \quad \nu(a) \leq a \leq \mu(a)]$$
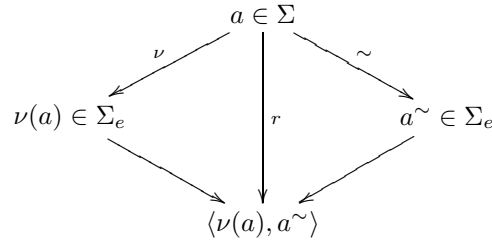
drawn in the following diagram:

Also in the case of BZ distributive lattice, $r(a)$ is the best approximation by exact elements of the element a. That is, for any element $a \in \Sigma$ there hold properties (L1)–(L3) and (U1)–(U3).

Moreover, in this context, the approximation r_e is given by the necessity–impossibility pair:

$$r_e(a) = \langle \nu(a), \mu'(a) \rangle = \langle \nu(a), a^\sim \rangle$$

drawn in the following diagram:

4.1. **Rough Approximation Spaces as BZ lattices.** In Section 3.1, we showed that it is possible to give to the collection of rough approximations induced by a quasi BZ distributive lattice the structure of quasi BZ distributive lattice. A dual result holds also in the case of BZ distributive lattices.

Let us consider the collection of all necessity–impossibility rough approximations. Analogously to the quasi BZ lattice case, we define

$$\forall a \in \Sigma \quad a_i := \nu(a) \quad a_e := a^\sim$$

so that the collection of rough approximations is $\mathbb{A} = \{\langle \nu(a), \sim a \rangle : a \in \Sigma\} = \{\langle a_i, a_e \rangle : a \in \Sigma\}$ with a_i and a_e exact sets: $a_i, a_e \in \Sigma_e$. Now, the following proposition holds for any BZ lattice.

Proposition 4.6. *The structure* $(\mathbb{A}, \sqcap, \sqcup, ^-, ^\approx, \langle 0,1 \rangle, \langle 1,0 \rangle)$ *where:*

$$\langle a_i, a_e \rangle \sqcap \langle b_i, b_e \rangle := \langle a_i \wedge b_i, a_e \vee b_e \rangle$$
$$\langle a_i, a_e \rangle \sqcup \langle b_i, b_e \rangle := \langle a_i \vee b_i, a_e \wedge b_e \rangle$$
$$\langle a_i, a_e \rangle^- := \langle a_e, a_i \rangle$$
$$\langle a_i, a_e \rangle^\approx := \langle a_e, (a_e)' \rangle$$

is a BZ^{dM} *distributive lattice.*

Also in this case the partial order induced by the lattice operators, $\sqcap$ and $\sqcup$, is

$$\langle a_i, a_e \rangle \sqsubseteq \langle b_i, b_e \rangle \quad \text{iff} \quad a_i \le b_i \text{ and } b_e \le a_e$$

In this environment the interior, closure and exterior operators are respectively defined as:

$$\mathbf{I}\langle a_i, a_e \rangle := \langle a_i, (a_i)' \rangle$$
$$\mathbf{C}\langle a_i, a_e \rangle := \langle (a_e)', a_e \rangle$$
$$\mathbf{E}\langle a_i, a_e \rangle := \langle a_e, (a_e)' \rangle$$

Obviously, the interior is equal to the necessity: $\mathbf{I}\langle a_i, a_e \rangle = \square \langle a_i, a_e \rangle$, and the closure is equal to the possibility: $\mathbf{C}\langle a_i, a_e \rangle = \Diamond \langle a_i, a_e \rangle$.

As can be seen, in terms of the original BZ lattice the modal operators of necessity $\square$ and possibility $\Diamond$ single out an exact element. That is, the necessity of a rough approximation $\langle a_i, a_e \rangle$ is equal to the rough approximation of the necessity of a:

$$(4.3) \qquad \square(r_e(a)) = \square\, \langle a_i, a_e \rangle = \langle a_i, (a_i)' \rangle = r_e(\nu(a))$$

Dually, the possibility of a rough approximation $\langle a_i, a_e \rangle$ is equal to the rough approximation of the possibility of a:

$$(4.4) \qquad \Diamond(r_e(a)) = \Diamond\, \langle a_i, a_e \rangle = \langle (a_e)', a_e \rangle = r_e(\mu(a))$$

The rough approximation of an element $\langle a_i, a_e \rangle$ is

$$\mathbf{R_e}\,\langle a_i, a_e \rangle = \langle \mathbf{I}\langle a_i, a_e \rangle, \mathbf{E}\langle a_i, a_e \rangle \rangle = \langle \langle a_i, (a_i)' \rangle, \langle a_e, (a_e)' \rangle \rangle$$

and according to Equations (17) and (18)

$$\mathbf{R_e}\,\langle a_i, a_e \rangle = \langle r_e(\nu(a)), r_e(\mu(a)) \rangle$$

4.2. **BZ lattices induced from Information Systems.** Now, we return to Pawlak rough sets

Proposition 4.7. *Let* $\langle X, Att(X), val(X), F \rangle$ *be a complete information system. Then, the structure* $\mathbb{P} = \langle \mathcal{P}(X), \cap, \cup, ^c, ^{\ne}, \emptyset, X \rangle$ *where, once fixed a set of attributes* $D \subseteq Att(X)$, *the preclusive relation* $^{\ne}$ *is defined*

$$H^{\ne} = \{x \in X : \forall y \in H\, \exists a \in D \quad F(x,a) \ne F(y,a)\}$$

is a BZ distributive lattice.

In general the structure $\mathbb{P}$ is not a de Morgan BZ lattice.

Example 4.8. Let us consider the flat information system of Table 1. We set $H = \{f_1, f_2, f_6\}$, $K = \{f_3, f_4, f_6\}$, $D = Att(X)$ and use the preclusive relation given by Equation (16). Then

$$(H \cap K)^\# = \{f_6\}^\# = \{f_1, f_2, f_3, f_4, f_5\} \neq \{f_2, f_4, f_5\} = \{f_4, f_5\} \cup \{f_2\} = H^\# \cup K^\#$$

∎

In this case, we have that the necessity (resp., possibility) operator corresponds to the "classical" rough sets lower (resp., upper) approximation:

$$\nu(H) = \mathbf{L}(H) = \{x \in X : H \subseteq E(x)\}$$

$$\mu(H) = \mathbf{U}(H) = \{x \in X : H \cap E(x) \neq \emptyset\}$$

where $E(x)$ is the granule generated by the element x.

So, given a set of objects $H \subseteq X$, its rough sets approximation is given by the pair $\langle \mathbf{L}(H), \mathbf{U}(H) \rangle$. Equivalently any rough set can be expressed as the lower - exterior pair $\langle \mathbf{L}(H), \mathbf{E}(H) \rangle = \langle \mathbf{L}(H), \mathbf{U}(H)^c \rangle$. By the previous subsection, we have that the collection of all pairs $\langle \mathbf{L}(H), \mathbf{E}(H) \rangle$ on an information system $\mathcal{K}(X) = \langle X, Att(X), val(X), F \rangle$, let us call it $\mathbb{RS}(X)$, gives rise to a BZ^{dM} distributive lattice $(\mathbb{RS}(X), \sqcap, \sqcup, {}^-, {}^\approx, \langle \emptyset, X \rangle, \langle X, \emptyset \rangle)$ where the operators are defined according to Proposition 4.6.

5. HW Algebras

In the previous section, we have seen that it is possible to give to Pawlak rough sets an algebraic representation through BZ^{dM} distributive lattices. However, this is not the only possibility. In literature, at least two other algebras are well known as a possible algebrization of rough sets. They are Heyting algebras and Wajsberg algebras ([Pag96, Pag98, Pol02]). In the following we will introduce Heyting Wajsberg algebras as a structure which admits as a model the collection of rough sets and as substructures all the above mentioned algebras ([CC02c]).

Definition 5.1. A system $\mathbf{A} = (A, \rightarrow_L, \rightarrow_G, 0)$ is a *Heyting Wajsberg (HW) algebra* if A is a non empty set, $0 \in A$, $\rightarrow_L, \rightarrow_G$ are binary operators, such that, once defined

(1) $a \vee b := (a \rightarrow_L b) \rightarrow_L b$
(2) $a \wedge b := ((a' \rightarrow_L b') \rightarrow_L b')'$
(3) $a' := a \rightarrow_L 0$
(4) $a^\sim := a \rightarrow_G 0$
(5) $1 := 0'$

the following are satisfied:

(HW1) $a \rightarrow_G a = 1$
(HW2) $(a \rightarrow_G b) \wedge b = b$
(HW3) $a \rightarrow_G (b \wedge c) = (a \rightarrow_G b) \wedge (a \rightarrow_G c)$
(HW4) $a \wedge (a \rightarrow_G b) = a \wedge b$
(HW5) $(a \vee b) \rightarrow_G c = (a \rightarrow_G c) \wedge (b \rightarrow_G c)$
(HW6) $1 \rightarrow_L a = a$
(HW7) $a \rightarrow_L (b \rightarrow_L c) = (a \rightarrow_L c)' \rightarrow_L b'$
(HW8) $a^{\sim\prime} \rightarrow_L a^{\sim\sim} = 1$
(HW9) $(a \rightarrow_G b) \rightarrow_L (a \rightarrow_L b) = 1$

It turns out that the operations $\wedge$ and $\vee$ are just the meet and join operators of a distributive lattice structure whose partial order is as usual defined as:

$$(5.1) \qquad a \le b \quad \text{iff} \quad a \wedge b = a$$

It is easy to prove that

$$(5.2a) \qquad a \le b \quad \text{iff} \quad a \to_L b = 1$$

$$(5.2b) \qquad \qquad \text{iff} \quad a \to_G b = 1$$

The primitive operator $\to_L$ behaves as a Łukasiewicz implication, as can be seen by the following proposition.

Proposition 5.2. *Let* **A** *be a HW algebra. Then the structure* $(A, \to_L, 1)$ *is a Wajsberg algebra according to* [Waj31, Waj35], *i.e., it satisfies:*

- (W1) $1 \to_L a = a$
- (W2) $(a \to_L b) \to_L ((b \to_L c) \to_L (a \to_L c)) = 1$
- (W3) $(a \to_L b) \to_L b = (b \to_L a) \to_L a$
- (W4) $(a' \to_L b') \to_L (b \to_L a) = 1$

Let us note that the structure of Wajsberg algebra induces a distributive lattice with respect to the partial order formally defined as in Equation (20a). In this lattice one can define the unary operator $a' := a \to_L 0$ which turns out to be a Kleene (or Zadeh) orthocomplementation. In other words properties (K1)–(K3) of Section 3 hold.

We remark that Wajsberg algebras are equivalent to Chang MV algebras ([Cha58]). Given a Wajsberg algerba, the MV disjunction $\oplus$ and conjunction $\odot$ operators are defined as:

$$(5.3a) \qquad a \oplus b := \neg a \to_L b$$

$$(5.3b) \qquad a \odot b := \neg (a \to_L \neg b)$$

Viceversa, given a MV algebra, the Łukasiewicz implication is defined as:

$$(5.4) \qquad a \to b := \neg a \oplus b$$

On the other hand, the primitive operator $\to_G$ behaves as a Gödel implication.

Proposition 5.3. *Let* **A** *be a HW algebra. Then by defining the lattice operators* $\wedge$ *and* $\vee$ *as in Definition (5.1), the partial order relation* $\le$ *as in Equation* (20b), *we have that* $(A, \to_G, \wedge, \vee,')$ *is a linear symmetric Heyting algebra according to* [Mon80], *i.e.,* **A** *satisfies the following properties (adopting the original Monteiro enumeration):*

- (A1) $a \to_G a = b \to_G b$
- (A2) $(a \to_G b) \wedge b = b$
- (A3) $a \to_G (b \wedge c) = (a \to_G b) \wedge (a \to_G c)$
- (A4) $a \wedge (a \to_G b) = a \wedge b$
- (A5) $(a \vee b) \to_G c = (a \to_G c) \wedge (b \to_G c)$
- (A6) $(a')' = a$
- (A7) $(a \wedge b)' = a' \vee b'$
- (K) $a \wedge a' \le b \vee b'$
- (L) $(a \to_G b) \vee (b \to_G a) = 1$

In particular, the unary operation $a^\sim = a \to_G 0$ is a *Brouwer* orthocomplementation. In other words properties (B1)–(B3) of Section 3 hold .

Finally, from the above propositions we have that any HW algebra induces a lattice structure endowed with a Kleene and a Brouwer negation, then it can be easily proved that HW algebras have as substructures distributive BZ^{dM} lattice.

Proposition 5.4. *Let* **A** *be a HW algebra. Then, by defining* $\wedge$ *and* $\vee$ *as in (5.1), we have that* $(A, \wedge, \vee, ', ^\sim, 0, 1)$ *is a distributive* BZ^{dM} *lattice.*

As a consequence of the latter proposition, we can define in any HW algebra a possibility (resp., necessity) operator and, as in BZ lattices, it behaves as a closure (resp., interior) operator. Further, these operators have an S_5 modal behaviour. So, in any HW algebra we can define a rough approximation space as in Definition 4.5. We recall that, given an element $a \in A$ its rough approximation is given by the necessity–possibility pair:

$$r(a) := \langle \nu(a), \mu(a) \rangle = \langle a'^\sim, a^{\sim\prime} \rangle \quad [\text{with} \quad \nu(a) \le a \le \mu(a)]$$

or equivalently, by the necessity–impossibility pair:

$$r_e(a) := \langle \nu(a), \mu'(a) \rangle = \langle a'^\sim, a^\sim \rangle$$

In the following we will see that it is possible to give the structure of a HW algebra to the collection of all rough approximations that can be obtained on a given BZ lattice, and consequently on a given HW algebra.

5.1. **Rough Approximation Spaces as HW algebras.** In Section 4.1, we have proved that it is possible to give to the collection of all rough approximations $\mathbb{A} := \{\langle \nu(a), a^\sim \rangle : a \in \Sigma\}$ on a BZ lattice $\langle \Sigma, \wedge, \vee, ', ^\sim, 0, 1 \rangle$ the structure of a BZ^{dM} lattice. We now prove that it holds the stronger result that $\mathbb{A}$ has the structure of a HW algebra.

Let us set, for the sake of simplicity $a_i = \nu(a)$ and $a_e = \mu(a)' = a^\sim$, and consider the collection of all necessity–impossibility pairs on a BZ distributive lattice $\mathbb{A} := \{\langle a_i, a_e : a \in \Sigma\}$. Then the following result holds:

Proposition 5.5. *The structure* $(\mathbb{A}, \Rightarrow_L, \Rightarrow_G, \mathbf{0})$, *where* $\Rightarrow_L, \Rightarrow_G$ *and* $\mathbf{0}$ *are defined as*

$$(5.5a) \qquad \langle a_i, a_e \rangle \Rightarrow_L \langle b_i, b_e \rangle := \langle ((a_i)' \wedge (b_e)') \vee a_e \vee b_i, a_i \wedge b_e \rangle$$

$$(5.5b) \qquad \langle a_i, a_e \rangle \Rightarrow_G \langle b_i, b_e \rangle := \langle ((a_i)' \wedge (b_e)') \vee a_e \vee b_i, (a_e)' \wedge b_e \rangle$$

$$(5.5c) \qquad\qquad\qquad \mathbf{0} := \langle 0, 1 \rangle$$

is a HW algebra.

According to Definition 5.1 it is possible to define the lattice operators on $\mathbb{A}$ as:

$$\langle a_i, a_e \rangle \sqcup \langle b_i, b_e \rangle := (\langle a_i, a_e \rangle \Rightarrow_L \langle b_i, b_e \rangle) \Rightarrow_L \langle b_i, b_e \rangle = \langle a_i \vee b_i, a_e \wedge b_e \rangle$$

$$\langle a_i, a_e \rangle \sqcap \langle b_i, b_e \rangle := ((\langle a_i, a_e \rangle^- \Rightarrow_L \langle b_i, b_e \rangle^-) \Rightarrow_L \langle b_i, b_e \rangle^-)^- = \langle a_i \wedge b_i, a_e \vee b_e \rangle$$

whose induced partial order is, as usual:

$$\langle a_i, a_e \rangle \sqsubseteq \langle b_i, b_e \rangle \quad \text{iff} \quad a_i \le b_i \text{ and } b_e \le a_e$$

According to Equations (21), the MV disjunction and conjunction operators are defined as:

$$\langle a_i, a_e \rangle \oplus \langle b_i, b_e \rangle := \langle (a_e^c \wedge b_e^c) \vee (a_i \vee b_i), a_e \wedge b_e \rangle$$

$$\langle a_i, a_e \rangle \odot \langle b_i, b_e \rangle := \langle a_i \wedge b_i, (a_i^c \wedge b_i^c) \vee (a_e \vee b_e) \rangle$$

The Kleene and Brouwer negations, the modal operators and the rough approximation are defined as in Subsection 4.1.

Of course, Theorem 5.5 holds for any BZ lattice. So, in particular, we can consider the structure $\mathbb{P} = \langle \mathcal{P}(X), \cap, \cup, {}^c, {}^{\neq}, \emptyset, X \rangle$ which is a BZ lattice thanks to Proposition 4.7. As a consequence, we have that the collection of Pawlak rough sets $\mathbb{RS}(X)$ on an information system $\mathcal{K}(X) = \langle X, Att(X), val(X), F \rangle$ has an HW algebraic structure $\langle \mathbb{RS}(X), \Rightarrow_L, \Rightarrow_G, \mathbf{0} \rangle$ once properly defined the operators $\Rightarrow_L$ and $\Rightarrow_G$ and the constant $\mathbf{0}$ according to Proposition 5.5.

Example 5.6. Let us consider the flats information system introduced in Table 1. Once defined the sets $H = \{f_1, f_4, f_5\}$, $K = \{f_2, f_4, f_6\}$ and considering all the attributes to define the equivalence relation $\mathcal{R}_D$, i.e. $D = Att(X)$, we have that the rough sets induced by H and K are respectively: $r_e(H) = \langle \{f_4, f_5\}, \{f_2, f_6\} \rangle$ and $r_e(K) = \langle \{f_2, f_6\}, \{f_1, f_3\} \rangle$. So the HW implications applied to $r_e(H)$ and $r_e(K)$ become:

$$\langle H_i, H_e \rangle \rightarrow_L \langle K_i, K_e \rangle = \langle \{f_2, f_6\}, \emptyset \rangle$$
$$\langle H_i, H_e \rangle \rightarrow_G \langle K_i, K_e \rangle = \langle \{f_2, f_6\}, \{f_1, f_3\} \rangle$$

Needless to stress $\emptyset \subseteq \{f_1, f_3\}$ so it is verified that

$$(\langle H_i, H_e \rangle \rightarrow_G \langle K_i, K_e \rangle) \sqsubseteq (\langle H_i, H_e \rangle \rightarrow_L \langle K_i, K_e \rangle)$$

The meet and join lattice operators are:

$$\langle H_i, H_e \rangle \sqcup \langle K_i, K_e \rangle = \langle \{f_2, f_4, f_5, f_6\}, \emptyset \rangle$$
$$\langle H_i, H_e \rangle \sqcap \langle K_i, K_e \rangle = \langle \emptyset, \{f_1, f_2, f_3, f_6\} \rangle$$

As can be seen, $r_e(H)$ and $r_e(K)$ are incomparable, i.e. neither $r_e(H) \sqsubseteq r_e(K)$ nor $r_e(K) \sqsubseteq r_e(H)$.
The MV disjunction and conjunction are respectively:

$$\langle H_i, H_e \rangle \oplus \langle K_i, K_e \rangle = \langle \{f_1, f_3, f_4, f_5\}, \emptyset \rangle$$
$$\langle H_i, H_e \rangle \odot \langle K_i, K_e \rangle = \langle \emptyset, \{f_1, f_2, f_3, f_6\} \rangle$$

Finally, the necessity and possibility of $r_e(H)$ are respectively:

$$\square \langle H_i, H_e \rangle = \langle \{f_4, f_5\}, \{f_1, f_2, f_3, f_6\} \rangle$$
$$\Diamond \langle H_i, H_e \rangle = \langle \{f_1, f_3, f_4, f_5\}, \{f_2, f_6\} \rangle$$

$\blacksquare$

6. Fuzzy Sets

In the previous sections we introduced some different algebraic approaches to the theory of rough sets. In particular, we have seen that it is possible to give a rich algebraic structure to the collection of rough sets on an information system. Now, we see how the same approach applies also to fuzzy sets by showing that the collection of fuzzy sets, once defined the suitable operators, has the structure of an HW algebra. This result leads us to the possibility of having a complete algebra which is able to fully characterize both rough and fuzzy sets.

Let us consider a set of objects X, called the Universe. A *fuzzy set* or *generalized characteristic functional* on X is defined as usual as a $[0, 1]$–valued function on X:

$$f : X \mapsto [0, 1]$$

We will indicate the collection of all fuzzy sets on X as $\mathcal{F}(X) = [0,1]^X$. This set contains two special elements: the identically 0 fuzzy set (also characteristic function of the empty set: $\underline{0} = \chi_\emptyset$)

$$\forall x \in X \quad \underline{0}(x) := 0$$

and the identically one fuzzy set (also characteristic function of the universe : $\underline{1} = \chi_X$)

$$\forall x \in X \quad \underline{1}(x) := 1$$

Obviously, this two elements will play the role of minimum and maximum element of the lattice structure we are going to define.

Proposition 6.1. *The structure $\langle [0,1]^X, \to_L, \to_G, \underline{0} \rangle$ is a HW algebra once defined the implication operators as follows*

$$(f_1 \to_L f_2)(x) := \min\{1, 1 - f_1(x) + f_2(x)\}$$

$$(f_1 \to_G f_2)(x) := \begin{cases} 1 & f_1(x) \leq f_2(x) \\ f_2(x) & otherwise \end{cases}$$

Example 6.2. Let us define the fuzzy set of *tall* and *very tall* people. So, we have that x is a real positive number, representing the height in centimeters. The tall fuzzy set is defined as:

$$t(x) := \begin{cases} 1 & x \geq 1.80 \\ \frac{x-1.50}{0.30} & 1.50 < x < 1.80 \\ 0 & x \leq 1.50 \end{cases}$$

and the very tall people fuzzy set is:

$$vt(x) := \begin{cases} 1 & x \geq 1.90 \\ \frac{x-1.60}{0.30} & 1.60 < x < 1.90 \\ 0 & x \leq 1.60 \end{cases}$$

If we compute the degree in which the fuzzy set very tall implies the fuzzy set tall, we correctly obtain with both implications the value 1:

$$vt \to_L t = \underline{1}$$
$$vt \to_G t = \underline{1}$$

that is, if one is very tall is surely tall. Of course, the viceversa is not true, and computing the implication of tall to very tall, we obtain:

$$(t \to_L vt)(x) := \begin{cases} 1 & x \leq 1.50 \quad or \quad x \geq 1.90 \\ \frac{1.80-x}{0.30} & 1.50 < x \leq 1.60 \\ \frac{2}{3} & 1.60 \leq x < 1.80 \\ \frac{x-1.60}{0.30} & 1.80 \leq x < 1.90 \end{cases}$$

in the case of $\to_L$ implication and

$$(t \to_G vt)(x) := \begin{cases} 1 & x \leq 1.50 \quad or \quad x \geq 1.90 \\ 0 & 1.50 < x \leq 1.60 \\ \frac{x-1.60}{0.30} & 1.60 \leq x < 1.90 \end{cases}$$

in the case of $\rightarrow_G$ operator. We remark that it holds an order relation among the two fuzzy sets $(t \rightarrow_G vt)$ and $(t \rightarrow_L vt)$: the first one is always less than or equal to the second one, as can be easily seen in Figure 6.2.
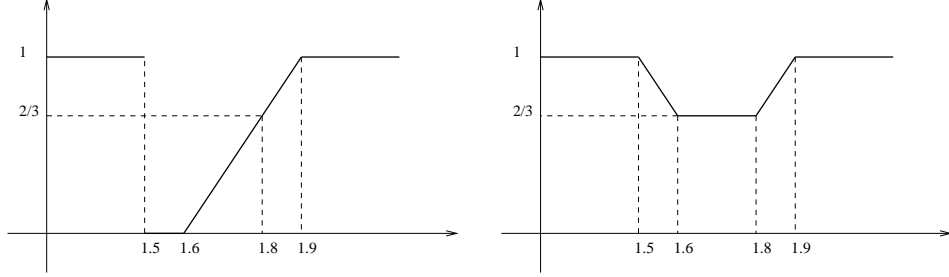

FIGURE 13. On the left it is drawn the set $(t \rightarrow_G vt)$ and on the right the set $(t \rightarrow_L vt)$.

∎

The lattice operators on $\mathcal{F}$ are defined as:

$$(f_1 \vee f_2)(x) := \max\{f_1(x), f_2(x)\}$$
$$(f_1 \wedge f_2)(x) := \min\{f_1(x), f_2(x)\}$$

The partial order relation induced on $\mathcal{F}$ by the above lattice operators is the usual pointwise ordering on fuzzy sets:

$$\forall f_1, f_2 \in [0,1]^X, \quad f_1 \leq f_2 \text{ iff } \forall x \in X, f_1(x) \leq f_2(x)$$

The MV operators of conjunction and disjunction are respectively defined as:

$$(f_1 \oplus f_2)(x) := \min\{1, f_1(x) + f_2(x)\}$$
$$(f_1 \odot f_2)(x) := \max\{0, f_1(x) + f_2(x) - 1\}$$

The Zadeh, Brouwer and anti–Brouwer negations are respectively defined as:

$$f'(x) := 1 - f(x)$$

$$f^\sim(x) := \begin{cases} 1, & \text{if } f(x) = 0 \\ 0, & \text{otherwise} \end{cases}$$

$$(\flat f)(x) := \begin{cases} 1 & f(x) \neq 1 \\ 0 & \text{otherwise} \end{cases}$$

For any fuzzy set $f \in [0,1]^X$ one can single out some peculiar subsets of the universe X:

• The *certainly-yes* domain of f:

$$\Delta_1(f) := \{x \in X : f(x) = 1\}$$

• The *certainly-no* domain of f:

$$\Delta_0(f) := \{x \in X : f(x) = 0\}$$

The necessity of a fuzzy set f is the characteristic functional of the certainly-yes domain of f:

$$\nu(f) = \chi_{\Delta_1(f)} = \begin{cases} 1 & \text{if } f(x) = 1 \\ 0 & \text{otherwise} \end{cases}$$

We can assume that two fuzzy sets *define the same property* iff they have the same certainly-yes domain. In this way, any property of fuzzy sets is associated with the certainly-yes domain $\Delta_1(f)$ of any of its elements. Such a property is interpreted as follows: *"belonging with certainty to the subset $\Delta_1(f)$ of X"*. Consequently any property is exactly represented by the characteristic functional $\chi_{\Delta_1(f)}$.

The possibility of a fuzzy set f is

$$\mu(f) = \begin{cases} 1 & \text{if } f(x) \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

Hence the impossibility operator is given by the characteristic functional of the certainly-no domain of f:

$$\mu'(f) = \chi_{\Delta_0(f)}$$

We can assume that two fuzzy sets *define the same noperty* iff they have the same certainly-no domain. Any noperty can be associated with the unique subset $\Delta_0(f)$ of X that represents the noperty in question: *"not belonging with certainty to the subset $\Delta_0(f)$ of X"*. As a consequence, our noperty is exactly represented by the characteristic functional $\chi_{\Delta_0(f)}$

References

[Cat97]   G. Cattaneo, *Generalized rough sets (preclusivity fuzzy-intuitionistic BZ lattices)*, Studia Logica **58** (1997), 47–77.

[Cat98]   _____, *Abstract approximation spaces for rough theories*, Rough Sets in Knowledge Discovery 1: Methodology and Applications (L. Polkowski and A.Skowron, eds.), Studies in Fuzziness and Soft Computing, Physica–Verlag, Heidelberg, 1998, pp. 59–98.

[CC02a]   G. Cattaneo and D. Ciucci, *A Quantative Analysis pf Preclusivity vs. Similarity Based Rough Approximations*, Lecture Notes in Artificial Intelligence, vol. 2475, Springer–Verlag, 2002, pp. 69–76.

[CC02b]   _____, *BZW algebras for an abstract approach to roughness and fuzziness*, IPMU 2002, July 1–5 2002, Annecy, France, Proceedings, ESIA – Université de Savoie, 2002, pp. 1103–1110.

[CC02c]   _____, *Heyting Wajsberg algebras as an abstract environment linking fuzzy and rough sets*, Lecture Notes in Artificial Intelligence, vol. 2475, Springer–Verlag, 2002, pp. 77–84.

[CCG98]   G. Cattaneo, M. L. Dalla Chiara, and R. Giuntini, *Some algebraic structures for many-valued logics*, Tatra Mountains Mathematical Publication **15** (1998), 173–196, Special Issue: Quantum Structures II, Dedicated to Gudrun Kalmbach.

[CGP99]   G. Cattaneo, R. Giuntini, and R. Pilla, *$BZMV^{dM}$ and Stonian MV algebras (applications to fuzzy sets and rough approximations)*, Fuzzy Sets Syst. **108** (1999), 201–222.

[Cha58]   C. C. Chang, *Algebraic analysis of many valued logics*, Trans. Amer. Math. Soc. **88** (1958), 467–490.

[Che88]   B. F. Chellas, *Modal logic, an introduction*, Cambridge University Press, Cambridge, MA, 1988.

[DO01]   I. Düntsch and E. Orlowska, *Beyond modalities: Sufficiency and mixed algebras*, Relational Methods for Computer Science Applications (E. Orlowska and A. Szalas, eds.), Physica–Verlag, Heidelberg, 2001, pp. 277–299.

[GMS98]   S. Greco, B. Matarazzo, and R. Slowinski, *Fuzzy similarity relation as a basis for rough approximations*, Rough Sets and Current Trends in Computing, First International Conference, RSCTC'98 (Warsaw) (Berlin) (L. Polkowski and A. Skowron, eds.), Lecture Notes in Computer Science, vol. 1424, Springer-Verlag, 1998, pp. 283–289.

[Kel55]   J.L. Kelley, *General topology*, Springer–Verlag, New York, 1955.

[Kry98]   M. Kryszkiewicz, *Rough set approach to incomplete information systems*, Information Sciences **112** (1998), 39–49.

[Mon80]   A. A. Monteiro, *Sur les algèbres de Heyting symétriques*, Portugaliae Mathematica **39** (1980), 1–237.

[NSS98]   S. Hoa Nguyen, A. Skowron, and P. Synak, *Discovery of data pattern with applications to decomposition and classification problems*, Rough Sets in Knowledge Discovery 2 (L. Polkowski and A. Skowron, eds.), Physica-Verlag, Heidelberg, 1998, pp. 55–97.

[Orl98a] E. Orlowska (ed.), *Incomplete information: Rough set analysis*, Physica–Verlag, Heidelberg, 1998.

[Orl98b] _____, *Introduction: What you always wanted to know about rough sets*, in *Incomplete Information: Rough Set Analysis* [Orl98a], pp. 1–20.

[Pag96]   P. Pagliani, *Rough sets and nelson algebras*, Fundamenta Informaticae **27** (1996), no. 2,3, 205–219.

[Pag98]   _____, *Rough set theory and logic–algebraic structures*, in Orlowska [Orl98a], pp. 109–190.

[Paw81]   Z. Pawlak, *Information systems - theoretical foundations*, Information Systems **6** (1981), 205–218.

[Paw82]   _____, *Rough sets*, Int. J. Inform. Comput. Sci. **11** (1982), 341–356.

[Paw85]   _____, *Rough sets and fuzzy sets*, Fuzzy Sets Syst. **17** (1985), 99–102.

[Paw92]   _____, *Rough sets: A new approach to vagueness*, Fuzzy Logic for the Management of Uncertainty (L. A. Zadeh and J. Kacprzyc, eds.), J. Wiley and Sons, New York, 1992, pp. 105–118.

[Paw98]   _____, *Rough set elements*, Rough Sets in Knowledge Discovery (L. Polkowski and A. Skowron, eds.), Studies in Fuzziness and Soft Computing Series, Physica-Verlag, Springer, Berlin, 1998, pp. 10–30.

[Pol02]   L. Polkowski, *Rough sets. mathematical foundations.*, Physica Verlag, Heidelberg, 2002.

[SS96]   A. Skowron and J. Stepaniuk, *Tolerance approximation spaces*, Fundamenta Informaticae **27** (1996), 245–253.

[ST99]   J. Stefanowki and A. Tsoukiàs, *On the extension of rough sets under incomplete information*, New Directions in Rough Sets, Data Mining, and Granular-Soft Computing, 7th International Workshop, RSFDGrC '99, Yamaguchi, Japan (Heidelberg) (Setsuo Ohsuga Ning Zhong, Andrzej Skowron, ed.), Lecture Notes in Computer Science, vol. 1711, Springer, 1999, pp. 73–81.

[ST01]   J. Stefanowski and A. Tsoukiàs, *Valued tolerance and decision rules*, Lecture Notes in Artificial Intelligence, vol. 2005, Springer-Verlag, 2001, pp. 212–219.

[Ste98]   J. Stepaniuk, *Approximation spaces in extensions of rough sets theory*, Rough Sets and Current Trends in Computing, First International Conference, RSCTC'98 (Warsaw, Poland) (Berlin) (L. Polkowski and A. Skowron, eds.), Lecture Notes in Computer Science, vol. 1424, Springer, 1998, pp. 290–297.

[Sur77]   S. Surma, *Logical works*, Polish Academy of Sciences, Wroclaw, 1977.

[SV95]   R. Słowinski and D. Vanderpooten, *Similarity relation as a basis for rough approximations*, Tech. Report 53, Institute of Computer Science, University of Technology, Warsaw, 1995.

[Vak91]   D. Vakarelov, *A modal logic for similarity relations in Pawlak knowledge representation systems*, Fundamenta Informaticae **XV** (1991), 61–79.

[Waj31]   M. Wajsberg, *Aksjomatyzacja trówartościowego rachunkuzdań [Axiomatization of the three-valued propositional calculus]*, Comptes Rendus des Séances de la Societé des Sciences et des Lettres de Varsovie **24** (1931), 126–148, English Translation in [Sur77].

[Waj35]   _____, *Beiträge zum Metaaussagenkalkül I*, Monashefte fur Mathematik un Physik **42** (1935), 221–242, English Translation in [Sur77].

Dipartimento Di Informatica, Sistemistica e Comunicazione. Università di Milano – Bicocca. Via Bicocca degli Arcimboldi 8, I–20126 Milano (Italia)
   E-mail address: {cattang,ciucci}@disco.unimib.it