*Fundamenta Informaticae XX (2006) 1–19*
*IOS Press*



# A possibility-theoretic view of formal concept analysis

**Didier Dubois**
*dubois@irit.fr*

**Florence Dupin de Saint-Cyr**
*bannay@irit.fr*

**Henri Prade**
*prade@irit.fr*

**Abstract.** The paper starts from the standard relational view linking objects and properties in formal concept analysis, here augmented with four modal-style operators (known as sufficiency, dual sufficiency, necessity and possibility operators). Formal concept analysis is mainly based on the first operator, while the others come from qualitative data analysis and can be also related to rough set theory. A possibility-theoretic reading of formal concept analysis with these four operators is proposed. First, it is shown that four and only four operators are indeed needed in order to describe the nine situations that can occur when comparing a statement (or its negation) with a state of information. The parallel between possibility theory and formal concept analysis suggests the introduction of new notions such as normalization and conditioning in the latter framework, also leading to point out some meaningful properties. Moreover, the graded setting of possibility theory allows us to suggest the extension of formal concept analysis to situations with incomplete or uncertain information.

## 1. Introduction

In the last three decades, a variety of basic frameworks aiming at processing different aspects of information, not really considered before, has blossomed in a series of works. Among them, possibility theory, rough set theory, and lattice-based formal concept analysis are the building blocks of various recent research trends in knowledge discovery. Indeed, in the late seventies, fuzzy sets have been used for developing a new theory of uncertainty representation, named possibility theory, which departs from probability theory by providing a proper setting for modeling incomplete information and which uses

Address for correspondence: IRIT, Université Paul Sabatier, 31062 Toulouse cedex 09 (France)

a disjunctive reading of the notion of set (distinct from the regular conjunctive interpretation used for ordinary sets and fuzzy sets)[25, 7, 10]. A few years later, the idea of rough sets has been advocated for acknowledging the fact that pieces of information can be expressed at different levels of granularity and that depending on this level, a set of objects may be only approximated from below and above in terms of classes of elements that are indiscernible (with respect to a set of properties used for describing them) [20, 21]. At about the same time, the duality between objects and properties has been exploited in a lattice theory setting already investigated before (e.g., [1]). It has led to an original view of the notion of a formal concept with an algorithmic concern [22, 14, 15].

These three theoretical frameworks have been developed independently until recently, although there exist works comparing rough sets with fuzzy sets and possibility theory, clarifying the differences and the possible cross-fertilizations between fuzzy sets and rough sets [8], rough sets with formal concept analysis [23] and investigating hybrid structures such as fuzzy formal concept analysis [2].

This paper is an attempt at comparing and combining possibility theory with formal concept analysis, in the same spirit as [8] for possibility theory and rough set theory. Indeed, we provide a unified set-theoretic view of the basic notions underlying formal concept analysis and possibility theory, starting with the relation objects/properties and its four modal-style set extensions (recalled in section 2). Section 3 justifies the need for four and only four operators. The main contributions of the paper come from the parallel that is drawn between the two fields in section 4. This leads to some new notions and related results described in section 5, and also to some generalizations of formal concept analysis when information may be incomplete or uncertain in section 6. Indeed, the existence of a link between an object and a property is usually assumed to be precisely known with complete certainty in rough sets or in formal concept analysis. Hence, in the last part of this paper we discuss the case where this link is incompletely known. Moreover, since possibility functions are naturally graded, this leads to a gradual extension of the four operators that handle the case of uncertain information.

## 2. Background on formal concept analysis setting

A formal information system is simply viewed here as a binary relation $R$ between a set $Obj$ of objects and a set $Prop$ of Boolean properties. $R$ is called *context* in formal concept analysis. If $X \subseteq Obj$, we denote by $\overline{X}$ its complementary set $Obj \setminus X$. We use the following notation $(x, y) \in R$ which means that object $x$ has property $y$. Let $R(x) = \{y \in Prop | (x, y) \in R\}$ be the set of properties of object $x$. Similarly we can define $R^{-1}(y) = \{x \in Obj | (x, y) \in R\}$, the set of objects that have property $y$.

**Example 2.1.** We consider an example of relation $R_0$ described by the table of Figure 1. This relation defines the links between eight objects $Obj = \{1, 2, 3, 4, 5, 6, 7, 8\}$ and nine properties $Prop = \{a, b, c, d, e, f, g, h, i\}$. There is a "$\times$" in the case corresponding to an object $x$ and to a property $y$ if the object $x$ has the property $y$, in other words the "$\times$"s describe the relation $R_0$ (or context). An empty case corresponds to the fact that $(x, y) \notin R_0$, i.e., it is known that object $x$ has not property $y$.

| | objects | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| properties | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| a | × | × | × | × | × | × | × | × |
| b | × | × | × | | × | × | | |
| c | | | × | × | | × | × | × |
| d | | | | | × | × | × | × |
| e | | | | | | | × | |
| f | | | | | × | × | | × |
| g | × | × | × | × | | | | |
| h | | × | × | × | | | | |
| i | | | | × | | | | |

Figure 1. $R_0$: a relation objects/properties borrowed from [15].

When $R(x)$ is extended to a subset $X$ of $Obj$, four remarkable sets can be defined:

$R^N(X) = \{y \in Prop | R^{-1}(y) \subseteq X\}$

$R^\Pi(X) = \{y \in Prop | R^{-1}(y) \cap X \neq \emptyset\}$

$R^\Delta(X) = \{y \in Prop | R^{-1}(y) \supseteq X\}$

$R^\nabla(X) = \{y \in Prop | R^{-1}(y) \cup X \neq Obj\}$

$R^N(X)$ is the set of properties such that any object that satisfies one of them is necessarily in $X$. In other words, each property of $R^N(X)$ is a sufficient condition for belonging to $X$.

$R^\Pi(X)$ can be rewritten as $\cup_{x \in X} R(x)$ and is the set of properties such that every object that satisfies one of them is possibly in $X$. In other words, if an object has no property $y$ that is in $R^\Pi(X)$ then it cannot belong to $X$. Indeed $R^N(X) = \overline{R^\Pi(\overline{X})} = Prop \setminus R^\Pi(\overline{X})$.

$R^\Delta(X)$ can be rewritten as $\cap_{x \in X} R(x)$ and is the set of properties in $Prop$ shared by all objects in $X$. In other words, satisfying all properties in $R^\Delta(X)$ is a necessary condition for an object to belong to $X$. $R^\Delta(X)$ is a partial conceptual characterization of objects in $X$: objects in $X$ should have all the properties of $R^\Delta(X)$ and may have some others (that are not shared by all objects in $X$). It is worth noticing that $\overline{R^\Pi(X)}$ provides a negative conceptual characterization of objects in $X$ since it gathers all the properties that are never satisfied by any object in $X$.

Note that $R^\nabla(X) = \overline{R^\Delta(\overline{X})} = Prop \setminus R^\Delta(\overline{X})$. Thus $R^\nabla$ is the set of properties in $Prop$ that are not satisfied by at least one object in $\overline{X}$.

**Example 2.2. (continued)**
If we consider the following subset of objects, here is the set of properties we obtain for the four defini-

tions:

| $X$ | $R_0^N(X)$ | $R_0^\Pi(X)$ | $R_0^\Delta(X)$ | $R_0^\nabla(X)$ |
|---|---|---|---|---|
| $\{1,2,3,4,5\}$ | $\{g,h,i\}$ | $\{a,b,c,d,f,g,h,i\}$ | $\{a\}$ | $\{b,e,f,g,h,i\}$ |
| $\{1,5\}$ | $\emptyset$ | $\{a,b,c,f,g\}$ | $\{a,b\}$ | $\{b,c,d,e,f,g,h,i\}$ |
| $\{2,3,4\}$ | $\{h,i\}$ | $\{a,b,c,g,h,i\}$ | $\{a,g,h\}$ | $\{b,c,d,e,f,g,h,i\}$ |
| $\{4\}$ | $\{i\}$ | $\{a,c,g,h,i\}$ | $\{a,c,g,h,i\}$ | $\{b,c,d,e,f,g,h,i\}$ |
| $\{1,4,5,6,7,8\}$ | $\{d,e,f,i\}$ | $\{a,b,c,d,e,f,g,h,i\}$ | $\{a\}$ | $\{c,d,e,f,i\}$ |

Figure 2 shows the four subsets of properties associated to the subset of objects $X = \{1,2,3,4,5\}$. As suggested in this figure, $R^\Delta(X)$ and $R^\Pi(X)$ respectively provide a lower bound and an upper bound of the set of properties that objects in $X$ may have: all objects in $\{1,2,3,4,5\}$ have property $a$ and none has property $e$. $R^\Delta(\overline{X})$ and $R^\Pi(\overline{X})$ provide similar information for the complementary set $\overline{X}$.

Note that the definitions giving a set of properties corresponding to a set of objects can be also easily modified for defining the set of objects relatively to a set of properties $Y \in Prop$: $R^{-1N}(Y)$, $R^{-1\Pi}(Y)$, $R^{-1\Delta}(Y)$, $R^{-1\nabla}(Y)$. Definitions and properties are similar for the dual side, and can be easily obtained by inverting $R$ and $R^{-1}$ and exchanging the role of the sets $Obj$ and $Prop$. However, in the following sections, we are going to consider only one side of the relation object/property.

In formal concept analysis, the pair of set valued functions $R^\Delta$ and $R^{-1\Delta}$ induces a Galois connection [4, 19] between $2^{Obj}$ and $2^{Prop}$. Then, a formal concept is a pair $(X, Y)$ where $X = R^{-1\Delta}(Y)$ and $Y = R^\Delta(X)$, $X$ is called its *extent* and $Y$ its *intent*. For instance, in our example, $(\{2,3,4\},\{a,g,h\})$ is a concept.

Besides, the four subsets $R^N(X)$, $R^\Pi(X)$, $R^\Delta(X)$, and $R^\nabla(X)$ have been considered in qualitative data analysis by Gediga and Düntsch [16, 11], where $R^\Delta$ is called *sufficiency* operator. Düntsch and Orlowska [12, 13] have studied the representation capabilities of sufficiency operators in the theory of Boolean algebras. In a rough set perspective, Yao [24, 23] has also considered the same four subsets (using the notations $X^\Box$, $X^\Diamond$, $X^*$, $X^\#$ respectively for $R^N(X)$, $R^\Pi(X)$, $R^\Delta(X)$, and $R^\nabla(X)$).

The definitions of $R^N$ and $R^\Pi$ look a bit like lower and upper approximation in rough set theory. However, in Pawlak's proposal[1] [20, 21], relation $R$ (defined here on $Obj \times Prop$) is replaced by an equivalence relation in $Obj$ among objects that are indiscernible because they satisfy the same set of properties. In [23], Yao shows how a rough set structure can be induced by reprojecting in the set of objects the result obtained by the application of $R^\Pi$ or $R^N$ namely, $R^{-1\Pi}(R^N(X)) \subseteq X \subseteq R^{-1N}(R^\Pi(X))$.

The reasons for our change of notations of the four operators will be made clear in the following when we shall establish a parallel of this framework with possibility theory and then extend it to incomplete information represented in this setting.

[1]If $E$ is an equivalence relation defined on $Obj \times Obj$ by $(x, x') \in E$ iff $R(x) = R(x')$, then the lower and the upper rough approximations of a subset $X \subseteq Obj$ are defined by $X_* = \{x|E(x) \subseteq X\}$ and $X^* = \{x|E(x) \cap X \neq \emptyset\}$ (clearly here $E = E^{-1}$). Note that the equivalence classes of $E$, which are supposed to gather the objects that are indiscernible w.r.t. a set of potential properties, are the largest subsets $X$ of $Obj$ such that $R^\Delta(X) = R^\Pi(X)$ in the framework discussed in this paper.

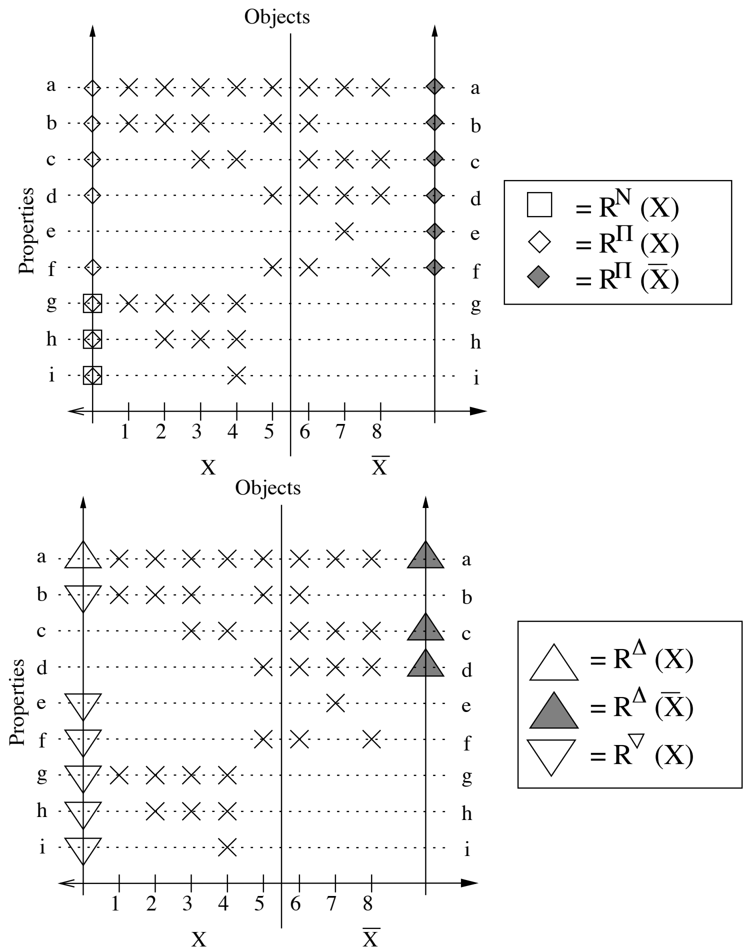


Figure 2. The four subsets of $Prop$ associated to $X = \{1, 2, 3, 4, 5\}$ w.r.t. the relation $R_0$.

| | | |
|---|---|---|
| 1 | A $\overline{A}$ | $E = \emptyset$ |
| 2 | E / A $\overline{A}$ | $E \neq \emptyset, E \neq U, A \subset \overline{E}$ |
| 3 | E / A $\overline{A}$ | $E \neq \emptyset, E \neq U, \overline{A} = E$ |
| 4 | E / A $\overline{A}$ | $E \neq \emptyset, E \neq U, \overline{A} \subset E$ |
| 5 | E / A $\overline{A}$ | $E \neq \emptyset, E \neq U, A \cap E \neq \emptyset, \overline{A} \cap E \neq \emptyset$ |
| 6 | E / A $\overline{A}$ | $E \neq \emptyset, E \neq U, E \subset A$ |
| 7 | E / A $\overline{A}$ | $E \neq \emptyset, E \neq U, A = E$ |
| 8 | E / A $\overline{A}$ | $E \neq \emptyset, E \neq U, A \subset E$ |
| 9 | E=U / A $\overline{A}$ | $E$ is equal to $U$ and $A$ is not equal to $U$ |

Figure 3. The nine situations of a set $A$ or its complement w.r.t. a given set $E$

## 3. Why four operators are needed?

By inspection of the formulas defining the four subsets $R^N(X)$, $R^\Pi(X)$, $R^\Delta(X)$, and $R^\nabla(X)$ we can see that their expressions involve Boolean relations (overlapping and inclusion) between the two sets $R^{-1}(y)$ and $X$. More generally, it can be checked (see Figure 3) that there are only nine mutually exclusive ways for describing the topological relations between a non-trivial set $A$ (i.e., $A \neq \emptyset$ and $A \neq U$) or its complement $\overline{A}$ (that implicitly refer to a request to evaluate) with respect to another given set $E$ (representing available information). Figure 3 enumerates all the possible situations starting from the case where $E$ is empty to the case where $E$ is the universe $U$ itself. Indeed this representation is faithful since it is possible to separate the elements of $A$ and $\overline{A}$ and to reorder $U$ in such a way that the elements of $E$ appear as a whole in the picture of Figure 3. The first four cases correspond to the situation where $E$ is empty, $E$ is strictly included in $\overline{A}$, $E$ is equal to $\overline{A}$ and $E$ contains $\overline{A}$ and overlaps on $A$ without including it. The fifth case is when $E$ both overlaps $A$ and $\overline{A}$ without including any of them. The last four cases correspond to the first four changing $E$ into $\overline{E}$.

We need four Boolean variables in order to represent these nine situations (three Boolean variables can cover only eight cases!). Let us consider the four following variables:

$\Pi(A) = 1$ if $A \cap E \neq \emptyset$, $\Pi(A) = 0$ otherwise
$\Delta(A) = 1$ if $A \subseteq E$, $\Delta(A) = 0$ otherwise
$\Pi(\overline{A}) = 1$ if $\overline{A} \cap E \neq \emptyset$, $\Pi(\overline{A}) = 0$ otherwise
$\Delta(\overline{A}) = 1$ if $\overline{A} \subseteq \emptyset$, $\Delta(\overline{A}) = 0$ otherwise

| Situation | $A \cap E \neq \emptyset$ $\Pi(A)$ | $A \subseteq E$ $\Delta(A)$ | $\overline{A} \cap E \neq \emptyset$ $\Pi(\overline{A})$ | $\overline{A} \subseteq E$ $\Delta(\overline{A})$ |
|---|---|---|---|---|
| 1. $E = \emptyset$ | 0 | 0 | 0 | 0 |
| 2. $E \neq \emptyset, E \neq U, A \subset \overline{E}$ | 0 | 0 | 1 | 0 |
| 3. $E \neq \emptyset, E \neq U, E = \overline{A}$ | 0 | 0 | 1 | 1 |
| 6. $E \neq \emptyset, E \neq U, E \subset A$ | 1 | 0 | 0 | 0 |
| 4. $E \neq \emptyset, E \neq U, \overline{A} \subset E$ | 1 | 0 | 1 | 1 |
| 5. $E \neq \emptyset, E \neq U, E \cap A \neq \emptyset, E \cap \overline{A} \neq \emptyset$ | 1 | 0 | 1 | 0 |
| 7. $E \neq \emptyset, E \neq U, E = A$ | 1 | 1 | 0 | 0 |
| 8. $E \neq \emptyset, E \neq U, A \subset E$ | 1 | 1 | 1 | 0 |
| 9. $E = U, A \neq U$ | 1 | 1 | 1 | 1 |

Figure 4. The nine possible tuples of values of $\Pi(A), \Delta(A), \Pi(\overline{A}), \Delta(\overline{A})$

Note that this definition implies the following constraints:

$$\Delta(A) \leq \Pi(A) \qquad \Delta(\overline{A}) \leq \Pi(\overline{A})$$

Moreover, we can see in the table of Figure 4 that there are only nine possible tuples of values of $\Pi(A), \Delta(A), \Pi(\overline{A}), \Delta(\overline{A})$ that are allowed by the two previous constraints. They correspond to the nine situations described above. These four Boolean variables are associated with four measures that have been proposed in possibility theory (see Annex), namely $\Pi$, $N$, $\Delta$ and $\nabla$ where $N(A) = 1 - \Pi(\overline{A})$ and $\nabla(A) = 1 - \Delta(\overline{A})$. They allow to characterize these nine cases.

# 4. Possibility-theoretic reading of formal concept analysis

In this section, we develop the parallel between possibility theory and the formal information systems framework, which has just been suggested in the previous section.

## 4.1. Relation-based possibility theory

Let the available information $E$ be the set of objects having property $y$, i.e., $E = R^{-1}(y)$, and let the set $A$ to evaluate be the set $X$ of objects. Then given a property $y \in Prop$, we can rewrite the four subsets introduced in formal concept analysis in a possibilistic manner as follows:

$$N_y(X) = \begin{cases} 1 & \text{if } R^{-1}(y) \subseteq X \\ 0 & \text{otherwise} \end{cases}$$

$$\Pi_y(X) = \begin{cases} 1 & \text{if } R^{-1}(y) \cap X \neq \emptyset \\ 0 & \text{otherwise} \end{cases}$$

$$\Delta_y(X) = \begin{cases} 1 & \text{if } R^{-1}(y) \supseteq X \\ 0 & \text{otherwise} \end{cases}$$

$$\nabla_y(X) = \begin{cases} 1 & \text{if } R^{-1}(y) \cup X \neq Obj \\ 0 & \text{otherwise} \end{cases}$$

We recognize the respective characteristic functions of the four sets $R^N(X)$, $R^\Pi(X)$, $R^\Delta(X)$ and $R^\nabla(X)$ (e.g., $y \in R^N(X)$ iff $N_y(X) = 1$ and similarly for the three other expressions). Note that $N_y(X) = 1 - \Pi_y(\overline{X})$ and $\nabla_y(X) = 1 - \Delta_y(\overline{X})$. They are also two-valued necessity, possibility, guaranteed possibility and potential necessity measures respectively, in the sense of possibility theory (see Annex).

As it is the case in possibility theory for finite settings, these four measures can be defined from a possibility distribution (which is the starting building block in this theory). Formally speaking, given a property $y \in Prop$, $R$ induces a two-valued possibility distribution $\pi_y$, such that:

$$\forall x \in Obj, \quad \pi_y(x) = \begin{cases} 1 & \text{if } (x, y) \in R \\ 0 & \text{otherwise,} \end{cases}$$

which is the characteristic function of $R^{-1}(y) = \{x \in Obj | (x, y) \in R\}$. Intuitively speaking, if all we know about an unknown object is that it has property $y$ then this object may be any $x$ such that $\pi_y(x) = 1$. Thus, the object/property relation $R$ (instantiated for a particular property $y$) in formal concept analysis plays the role of the possibility distribution $\pi$ encoding a standard granule of information "V is E" in possibility theory (see Annex).

As it can be checked, $N_y(X)$, $\Pi_y(X)$, $\Delta_y(X)$ and $\nabla_y(X)$ are respectively the necessity measure, the possibility measure, the guaranteed possibility measure and the potential certainty measure based on $\pi_y$ in the sense of possibility theory (see Annex for a refresher on these notions). Namely,

$$\begin{aligned} \Pi_y(X) &= \max_{x \in X} \pi_y(x) \\ N_y(X) &= \min_{x \notin X} (1 - \pi_y(x)) \\ \Delta_y(X) &= \min_{x \in X} \pi_y(x) \\ \nabla_y(X) &= \max_{x \notin X} (1 - \pi_y(x)) \end{aligned}$$

## 4.2. Object/property normalizations and noticeable inclusions

In the possibility framework theory (recalled in Annex) the normalization of a possibility distribution is required if we want to guarantee that for any event $A$, $\Pi(A) \geq N(A)$. Moreover, the normalization condition impossibility expresses the natural requirement that at least one value $u$ in the universe of discourse $U$ is completely possible. The counterpart to this condition now reads that any property is held by at least one object (otherwise this property is useless).

**Definition 4.1. (nn-normalization)**

$R^{-1}$ is not-none-normalized (nn-normalized for short) if $\forall y \in Prop, R^{-1}(y) \neq \emptyset$

As said before, this definition can be dually stated for enforcing that any object holds at least one property ($R$ is nn-normalized if $\quad \forall x \in Obj, \quad R(x) \neq \emptyset$).

**Proposition 4.1.** If $R$ is nn-normalized then $\forall X \subseteq Obj, \quad X \neq \emptyset \Rightarrow R^{\Pi}(X) \neq \emptyset$

This proposition comes from the fact that $R^{\Pi}(X) = \cup_{x \in X} R(x)$. Note that $R$ maybe nn-normalized without having $R^{-1}$ nn-normalized and conversely. Under the $R^{-1}$ nn-normalization condition we have the following inclusion:

**Proposition 4.2.** If $R^{-1}$ is nn-normalized then $\forall X \subseteq Obj, \quad R^{N}(X) \subseteq R^{\Pi}(X)$

Note that the property may not hold when $R^{-1}$ is not nn-normalized. For instance if we consider a property $y_0$ such that $R^{-1}(y_0) = \emptyset$ then $\forall X \subseteq Obj, y_0 \in R^{N}(X)$ and $y_0 \notin R^{\Pi}(X)$. On example 2.1, we can see that $R_0$ and $R_0^{-1}$ are nn-normalized, indeed there is no empty line nor empty column. We can also check that the property is verified on all the subsets of objects given in example 2.2.

Similarly, the counterpart of the normalization of $1-\pi$ expresses that any property should not be held by every object (otherwise this property is not very interesting in order to discriminate between subsets of objects).

**Definition 4.2. (na-normalization)**

$R^{-1}$ is not-all-normalized (na-normalized for short) if $\forall y \in Prop, \quad R^{-1}(y) \neq Obj$

Under this condition we have the following inclusion:

**Proposition 4.3.** If $R^{-1}$ is na-normalized then $\forall X \subseteq Obj, \quad R^{\Delta}(X) \subseteq R^{\nabla}(X)$

The reader can check that this inclusion does not hold in the example 2.1 due to the failure of the na-normalization of $R_0^{-1}$ since property $a$ is possessed by all objects (there is a full line of $\times$ in the matrix). Note that if we remove the property $a$ then the proposition holds (it can be checked for the five subsets given in the table of example 2.2 (forgetting $a$)).

Besides, without any normalization condition we always have:

**Proposition 4.4.** $R^{\Delta}(X) \subseteq R^{\Pi}(X)$ and $R^{N}(X) \subseteq R^{\nabla}(X)$

Then putting all these inclusion relations together, we obtain

**Proposition 4.5.** If $R^{-1}$ is na-normalized and nn-normalized then

$$\forall X \subseteq Obj, \quad R^{N}(X) \cup R^{\Delta}(X) \subseteq R^{\Pi}(X) \cap R^{\nabla}(X)$$

This is the counterpart of proposition 7.1 in possibility theory (see Annex).

**Remark 4.1.** If $X$ is a singleton then $R^{\Pi}(X) = R^{\Delta}(X)$. Lastly, if $\forall y \in Prop, \exists! x \in Obj$ such that $(x, y) \in R$ then $R^{N}(X) = R^{\Pi}(X)$.

## 5. Decomposability, conditioning and orthogonality

If we consider two subsets of objects $X_1$ and $X_2$, it is worth saying what holds for the four subsets of properties associated with $X_1 \cup X_2$ and with $X_1 \cap X_2$. This is done in the following proposition.

**Proposition 5.1.** (decomposability) $\forall X_1, X_2 \in Obj$,

$$
\begin{array}{ll}
R^N(X_1 \cup X_2) \supseteq R^N(X_1) \cup R^N(X_2) & R^N(X_1 \cap X_2) = R^N(X_1) \cap R^N(X_2) \\
R^\Pi(X_1 \cup X_2) = R^\Pi(X_1) \cup R^\Pi(X_2) & R^\Pi(X_1 \cap X_2) \subseteq R^\Pi(X_1) \cap R^\Pi(X_2) \\
R^\Delta(X_1 \cup X_2) = R^\Delta(X_1) \cap R^\Delta(X_2) & R^\Delta(X_1 \cap X_2) \supseteq R^\Delta(X_1) \cup R^\Delta(X_2) \\
R^\nabla(X_1 \cup X_2) \subseteq R^\nabla(X_1) \cap R^\nabla(X_2) & R^\nabla(X_1 \cap X_2) = R^\nabla(X_1) \cup R^\nabla(X_2)
\end{array}
$$

On the example 2.2, it can be checked for instance that $R_0^N(\{1,2,3,4,5\}) = \{g,h,i\}$ is indeed only a superset of the union of $R_0^N(\{2,3,4\}) = \{h,i\}$ with $R_0^N(\{1,4,5\}) = \{i\}$, while $R_0^\nabla(\{4\}) = \{b,c,d,e,f,g,h,i\}$ is indeed equal to the union of $R_0^\nabla(\{1,2,3,4,5\}) = \{g,h,i\}$ with $R_0^\nabla(\{2,3,4\}) = \{b,c,d,e,f,g,h,i\}$.

Conditioning is an important notion in possibility theory as recalled in the Annex. Its counterpart in the relational setting of formal concept analysis is defined by looking for the greatest solution of the equation below when $R^\Pi(X_1) \neq \emptyset$:

$$R^\Pi(X_1 \cap X_2) = R^\Pi(X_2|X_1) \cap R^\Pi(X_1)$$

with

$$R^\Pi(X_2|X_1) = \begin{cases} Prop & \text{if} \quad R^\Pi(X_1 \cap X_2) = R^\Pi(X_1) \neq \emptyset \\ R^\Pi(X_1 \cap X_2) & \text{if} \quad R^\Pi(X_1 \cap X_2) \subset R^\Pi(X_1) \end{cases}$$

$R^\Pi(X_2|X_1)$ is indeed the largest set of properties such that the above equation holds. Thus, the Π-conditioning here sanctions the fact that in the context of the set of objects in $X_1$, focusing on the subset $X_1 \cap X_2$ may reduce the set of possible properties for the considered objects. In case $R^\Pi(X_1 \cap X_2) \subset R^\Pi(X_1)$, there are less properties that are possible for objects in $X_1 \cap X_2$ than for objects in $X_1$. This means that the objects of $X_2$ in the context $X_1$ are not fully representative (typical) of the variety of properties of objects in $X_1$. We can define $R^N(X_2|X_1)$ by: $R^N(X_2|X_1) = \overline{R^\Pi(\overline{X_2}|X_1)}$. We have

$$R^N(X_2|X_1) \neq \emptyset \quad \Leftrightarrow \quad R^\Pi(X_1 \cap X_2) \supset R^\Pi(X_1 \cap \overline{X_2}),$$

$R^N(X_2|X_1) \neq \emptyset$ expresses that more properties become possible when going from $\overline{X_2}$ to $X_2$ in the "context" $X_1$.

Similarly, by analogy with the situation in standard possibility theory (see Annex)

$$R^\Delta(X_1 \cap X_2) = R^\Delta(X_2|X_1) \cup R^\Delta(X_1)$$

with

$$R^\Delta(X_2|X_1) = \begin{cases} \emptyset & \text{if} \quad R^\Delta(X_1 \cap X_2) = R^\Delta(X_1) \neq Prop \\ R^\Delta(X_1 \cap X_2) & \text{if} \quad R^\Delta(X_1 \cap X_2) \supset R^\Delta(X_1) \end{cases}$$

$R^{\Delta}(X_2|X_1)$ is the smallest set of properties such that the above equation holds. We have

$$R^{\Delta}(X_2|X_1) \neq \emptyset \Leftrightarrow R^{\Delta}(X_1 \cap X_2) \supset R^{\Delta}(X_1 \cap \overline{X_2})$$

i.e., going from $\overline{X_2}$ to $X_2$ in the context $X_1$, more properties become certain.

Only the inclusion $R^{\Pi}(X_1 \cap X_2) \subseteq R^{\Pi}(X_1) \cap R^{\Pi}(X_2)$ holds in general. Indeed, there may exist properties that are both in $R^{\Pi}(X_1)$ and in $R^{\Pi}(X_2)$, without being possible for those objects in $X_1 \cap X_2$ (because they are only possible in $X_1 \cap \overline{X_2}$ and in $\overline{X_1} \cap X_2$). The equality

$$R^{\Pi}(X_1 \cap X_2) = R^{\Pi}(X_1) \cap R^{\Pi}(X_2)$$

holds in situations, called "unrelatedness" by Nahmias [18] in standard possibility theory. Note that

$$R^{\Pi}(X_1 \cap X_2) = R^{\Pi}(X_1) \cap R^{\Pi}(X_2) \Leftrightarrow \begin{cases} R^{\Pi}(X_1 \cap \overline{X_2}) \subseteq R^{\Pi}(X_1 \cap X_2) \\ \text{or} \\ R^{\Pi}(\overline{X_1} \cap X_2) \subseteq R^{\Pi}(X_1 \cap X_2) \end{cases}$$

It means that possible properties for objects in $X_1 \cap X_2$ are equal to the possible properties for objects in $X_1$ that are possible for objects in $X_2$ if and only if
- either objects in $X_1$ that are not in $X_2$ do not have more possible properties than objects belonging both to $X_1$ and $X_2$
- or objects in $X_2$ that are not in $X_1$ do not have more possible properties than objects belonging both to $X_1$ and $X_2$.

Similarly, when the equality $R^{\Delta}(X_1 \cap X_2) = R^{\Delta}(X_1) \cap R^{\Delta}(X_2)$ holds, it means that the properties that for sure are possessed by the objects in $X_1 \cap X_2$ (or if we prefer that are common to all objects in $X_1 \cap X_2$ are only the properties that are common to all objects in $X_1$ and the properties that are common to all objects in $X_2$.

$\Pi$-unrelatedness and $\Delta$-unrelatedness are forms of *independence* between subsets of objects. Independence is a key concept in artificial intelligence, having well defined notions may allow to parallelize the handling of independent objects, to avoid combinatorial explosion when modeling a system (like in the frame problem [17]) etc. Other possible definitions of independence exist. Standard possibility theory distinguishes between weak ($\Pi(A \cap B) = \min(\Pi(A), \Pi(B))$) and strong ($N(A|B) = N(A)$) forms of independence (see [6]). Strong independence implies weak independence in standard possibility theory. A question beyond this introductory paper is the study of all the counterparts of these notions in the setting of formal concept analysis. In particular, it would be interesting to check if the implication between strong and weak independence still holds, and to study the links between the strong independence, and the notion of reduct in rough set theory [21].

The related notion of orthogonality between sets of objects or sets of properties is also worth of interest. Intuitively we mean that two subsets of objects $X_1$ and $X_2$ are orthogonal if they have no common properties. This writes $R^{\Pi}(X_1) \cap R^{\Pi}(X_2) = \emptyset$ if we mean "no common possible properties", or $R^{\Delta}(X_1) \cap R^{\Delta}(X_2) = \emptyset$ if we restrict to the properties shared in each subset of objects. This leads finally to four possible definitions of orthogonality:

**Definition 5.1.** (orthogonality)

$$X_1 \text{ is } \Pi\text{-orthogonal to } X_2 \quad \text{iff} \quad R^\Pi(X_1) \cap R^\Pi(X_2) = \emptyset$$

The definitions of $N$, $\Delta$ and $\nabla$-orthogonality are identical except that $R^\Pi$ is replaced respectively by $R^N$, $R^\Delta$ and $R^\nabla$.

**Proposition 5.2.** $R^\Pi(X_1) \cap R^\Pi(X_2) = \emptyset \Rightarrow X_1 \cap X_2 = \emptyset$ in case of nn-normalization of $R$

This proposition follows from proposition 5.1 and from proposition 4.1, indeed, when $R$ is nn-normalized, any object has at least one property.

$\Delta$-orthogonality is weaker and can equivalently be written $R^\Delta(X_1 \cup X_2) = \emptyset$ expressing that objects in $X_1 \cup X_2$ have nothing in common.

**Proposition 5.3.** $\Pi$-orthogonality implies $\Delta$-orthogonality, $\nabla$-orthogonality implies $N$-orthogonality
If $R^{-1}$ is nn-normalized, then $\Pi$-orthogonality implies $N$-orthogonality
If $R^{-1}$ is na-normalized, then $\nabla$-orthogonality implies $\Delta$-orthogonality

This proposition comes from propositions 4.2, 4.3, 4.4.

# 6. Further research: incompleteness and gradual uncertainty

In the previous framework, two important assumptions were made:

- we have complete information about the stated links between properties in $Prop$ and objects in $Obj$. Namely, $(x, y) \in R$ means that object $x$ satisfies property $y$ and $(x, y) \notin R$ means that object $x$ does not satisfy property $y$, rather than "we do not know if $(x, y) \in R$". Moreover, all the existing links are stated.
- the properties are supposed to be Boolean. Hence, when an object satisfies a property, it fully satisfies it: there is no intermediary degree of satisfaction since the property is not gradual.

Clearly each of the above assumptions may be relaxed. Namely, one may consider that there are pairs $(x, y)$ for which it is not known at all if $x$ has property $y$ or not. We may also have uncertain information such that we are certain at level $\alpha$ that $x$ satisfies $y$, or at level $\beta$ that $x$ does not satisfy $y$.

When relaxing the second assumption, we may still assume that we have complete information, but now about the satisfaction of non-Boolean properties. Namely, given a gradual property $y$, $\mu_{R^{-1}(y)}(x) = \alpha$ would then denote the fact that object $x$ satisfies property $y$ at degree $\alpha$ where $\mu_{R^{-1}(y)}$ is the membership function of the fuzzy set of objects that constitutes the extension of $R^{-1}(y)$. Such an extension has been considered by Belohlavek [2], see also [5].

In the following, we are only considering the first extension, where there are pairs $(x, y)$ such that one is uncertain if property $y$ applies or not to object $x$. Combining the two extensions would also make sense (when one may not be sure that some property $y$ is possessed by an object $x$ at least at some degree $\alpha$). However, it can not be developed here.

In the classical setting there are only two possible situations for any pair $(x, y)$: either $(x, y) \in R$ or $(x, y) \notin R$. In case of incomplete or uncertain information, we may distinguish between the same five typical situations as done in Annex w.r.t. the information $(x, y) \in R$, namely,

1. we are fully certain that $(x, y) \in R$
2. we are $\alpha$-certain that $(x, y) \in R$ (with $0 < \alpha < 1$) and thus it is $1 - \alpha$-possible that $(x, y) \notin R$
3. we are in a situation of *complete ignorance*, both $(x, y) \in R$ and $(x, y) \notin R$ are fully possible
4. we are $\beta$-certain that $(x, y) \notin R$ (with $0 < \alpha < 1$) and thus it is $1 - \beta$-possible that $(x, y) \in R$ (it means that it is all the less possible that $(x, y) \in R$ as $\beta$ is larger)
5. we are fully certain that $(x, y) \notin R$.

In case information is just incomplete, which means that in the cells $(x, y)$ of the table, we have '+' if $(x, y) \in R$, '-' if $(x, y) \notin R$, and nothing if we do not know, only situations 1, 3 and 5 exist, or if we prefer $\alpha = 1 = \beta$.

Then the set $R^{-1}(y)$ of objects that have property $y$ is now approximated from above by the fuzzy set $R^{-1\Pi}(y)$ of objects that have *possibly* property $y$ and by the fuzzy set $R^{-1N}(y)$ of objects that have *certainly* the property $y$, with the respective membership functions $\mu_{R^{-1\Pi}(y)}$ and $\mu_{R^{-1N}(y)}$. In case information is incomplete but certain, $R^{-1\Pi}(y) = \{x|$ it is not known that $(x, y) \notin R\}$ and $R^{-1N}(y) = \{x|$ it is known that $(x, y) \in R\}$. Clearly, $R^{-1N}(y) \subseteq R^{-1\Pi}(y)$. Let $(\gamma, \beta)_{(x,y)}$ be the pair of necessity (certainty) degrees that are now given in each cell $(x, y)$ of the object/property table in the general case. As already said, there are five situations $(1, 0)$, $(\gamma, 0)$, $(0, 0)$, $(0, \beta)$ and $(0, 1)$ corresponding respectively to full certainty and $\gamma$-certainty that $(x, y) \in R$, complete ignorance, $1 - \beta$ possibility that $(x, y) \in R$ and impossibility of $(x, y) \in R$. Assume $\min(\gamma, \delta) = 0$, or equivalently $\max(1 - \gamma, 1 - \delta) = 1$, which expresses normalization (at least $(x, y) \in R$ or $(x, y) \notin R$ is completely possible), then

$$\mu_{R^{-1\Pi}(y)}(x) = 1 - \delta \quad \text{and} \quad \mu_{R^{-1N}(y)}(x) = \gamma$$

Note that

$$\mu_{R^{-1N}(y)}(x) = \gamma > 0 \Rightarrow \mu_{R^{-1\Pi}(y)}(x) = 1$$

and

$$\mu_{R^{-1\Pi}(y)}(x) = 1 - \delta < 1 \Rightarrow \mu_{R^{-1N}(y)}(x) = 0$$

Thus $\forall y \in Prop$, the fuzzy set inclusion $R^{-1}(y) \subseteq R^{-1\Pi}(y)$ holds. Moreover, the membership functions of the fuzzy sets of properties that are possibly (resp. necessarily) possessed by some objects in $X$ is given by:

$$N_y(X) = \min_{x \notin X} 1 - \mu_{R^{-1\Pi}(y)}(x)$$
$$\Pi_y(X) = \max_{x \in X} \mu_{R^{-1\Pi}(y)}(x)$$

$\Delta_y(X)$ is the membership degree of $y$ to the set of properties certainly shared by the objects in $X$.

$$\Delta_y(X) = \min_{x \in X} \mu_{R^{-1N}(y)}(x)$$
$$\nabla_y(X) = \max_{x \notin X} 1 - \mu_{R^{-1N}(y)}(x)$$

The reason why $R^{-1N}(y)$ is used instead of $R^{-1\Pi}(y)$ in $\Delta$ (and $\nabla$), is due to the fact that $\Delta_y(X)$ should remain the set of properties that any object in $X$ more or less *certainly* has. It can be checked that decomposability properties described in proposition5.1 are preserved. In case of complete and certain information, $R^{-1\Pi}(y) = R^{-1N}(y) = R^{-1}(y)$, and the setting of section 4 is retrieved. In case information is just incomplete but not uncertain, the four modal-style operators of the formal concept analysis setting, are defined respectively by $R^{\Pi}(X) = \{y \in Prop | R^{-1\Pi}(y) \cap X \neq \emptyset\}$,$R^{N}(X) = \{y \in Prop | R^{-1\Pi}(y) \subseteq X\}$, $R^{\Delta}(X) = \{y \in Prop | R^{-1N}(y) \supseteq X\}$ and $R^{\nabla}(X) = \{y \in Prop | R^{-1N}(y) \cup X \neq Obj\}$.

## 7. Conclusion

This paper has only outlined the parallel between possibility theory and formal concept analysis and the associated modal-style operators. Clearly, much has still to be done for completely taking advantage of the parallel for discussing notions such as independency, Galois connection, orthogonality or redundancy.

Besides this parallel may be also fruitful for extending algorithms of formal concept analysis and rough set literature to the cases where information is incomplete or uncertain, or when properties may be graded. Moreover, possibility theory has a logical counterpart - possibilistic logic - where the four modal operators can be handled, while modal logics have been developed between the two settings in the rough set and formal concept analysis settings. This suggests a systematic study of the parallel at the logical and inference levels.

## Annex: Background on possibility theory

Zadeh [25] introduced possibility theory as a new framework for representing imprecise information, especially the one provided in linguistic terms. Thus, a piece of information of the form "$V$ is $E$", understood as "the possible values of the (single-valued) variable $V$ (supposed to range on a universe $U$) are restricted by a subset $E$ of $U$ (which may be fuzzy)". In other words, given the granule of information "$V$ is $E$", we know that the value of $V$ must/should be in $E$". Since $E$ may be fuzzy, this is understood as a graded possibility assignment, namely

$$\forall u \in U, \pi(u) = \mu_E(u)$$

where $\pi$ is a *possibility distribution*, defined as a function from $U$ to $[0, 1]$. It is important here to have in mind that the set $E$ should be understood disjunctively, in the sense that the variable $V$ is single-valued, and that the values restricted by $E$ (that belongs to $E$ with a non-zero degree) are mutually exclusive as possible values of $V$. The following conventions are assumed : i) $\pi(u) = 0$ means that $V = u$ is impossible, totally excluded ; ii) $\pi(u) = 1$ means that $V = u$ is fully possible, but several distinct $u$ and $u'$ are allowed to be simultaneously such that $\pi(u) = 1$ and $\pi(u') = 1$ ; iii) the larger $\pi(u)$, the more possible, i.e., plausible or feasible $u$ is. Thus, a possibility distribution encodes a *flexible* restriction (constraint). The above equality $\pi(u) = \mu_E(u)$ states that $V = u$ is possible, inasmuch as $u$ is compatible with $E$, given the piece of information "$V$ is $E$". The compatibility of $u$ with $E$ is estimated as the degree of membership function $\mu_E(u)$ of $u$ to $E$. Two more remarks are of importance here. First, the range $[0, 1]$ of $\pi$ may be replaced by any linearly ordered scale, possibly finite. Second, although in the examples originally used by Zadeh, and many authors after him, $U$ is a continuum ($U$ is a subset of the real line), $U$ may be finite and is not necessarily ordered. Thus, $U$ may be a set of possible states of the world that are encoded by the interpretations associated with a classical propositional logic language, and then $\pi$ rank-orders the possible states by their level of possibility according to the available information. In the following, we only consider the case where the universe $U$ is finite.

Possibility theory is a framework for modeling incomplete information about the (unique) value of some variable of interest. Since incomplete information is characterized by the fact that several values may remain more or less possible for the variable, the possibility distribution restricts a set of mutually exclusive values for the variables, leading to a disjunctive view of (fuzzy) sets.

## Possibility and necessity measures

Two set functions, also called "measures", are associated with a possibility distribution $\pi$, namely

- a possibility measure $\Pi$ (or "potential possibility") defined by

$$\Pi(A) = \max_{u \in A} \pi(u),$$

  which estimates to what extent the classical event $A$ is *consistent with* the information "$V$ is $E$" represented by $\pi$ (note that by definition $\Pi(\emptyset) = 0$);

- a dual measure of necessity $N$, expressing that an event is all the more necessarily (certainly) true as the opposite event is more impossible. $N$ is defined by

$$N(A) = 1 - \Pi(\overline{A}) = 1 - \max_{u \notin A} \pi(u),$$

  where $\overline{A} = U \backslash A$, is the complement of $A$. $N(A)$ estimates to what extent event $A$ is *implied* by the information "$V$ is $E$" represented by $\pi$ (inasmuch as this information entails that any realization of $\overline{A}$ is more or less impossible). Note that $N(U) = 1$ (since $\Pi(\emptyset) = 0$).

These two measures extend to fuzzy events. Then the degree of consistency of $A$ and $E$, is defined by

$$\Pi(A) = \max_{u \in U} \min(\mu_A(u), \pi(u)),$$

where $\mu_A$ is the membership function of $A$.

As a consequence, the possibility and necessity measures $\Pi$ and $N$ satisfy the characteristic properties (for classical or fuzzy events)

$$\Pi(A \cup B) = \max(\Pi(A), \Pi(B)) \text{ and } N(A \cap B) = \min(N(A), N(B)).$$

Note that we only have the inequalities:

$$\Pi(A \cap B) \leq \min(\Pi(A), \Pi(B)) \text{ and } N(A \cup B) \geq \max(N(A), N(B)).$$

Indeed, when $B$ is equal to $\overline{A}$ the left-hand side of the inequalities are respectively to $\Pi(\emptyset) = 0$ and $N(U) = 1$ while the right-hand side may take non extreme values. In case of total ignorance ($\forall u \in U, \pi(u) = 1$) we have $\Pi(A) = \Pi(\overline{A}) = 1$. The last inequality expresses that we may be somewhat certain that the true state of the world is in $A \cup B$ without being sure if it is in $A$ or if it is in $B$.

Assuming that the possibility distribution $\pi$ is normalized ($\exists u \in U, \pi(u) = 1$) amounts to expressing that there is no underlying contradiction pervading the available information (at least one value of $U$ is completely possible). The normalization of $\pi$ ensures that $\forall A, \Pi(A) \geq N(A)$, i.e., any event $A$ should be possible before being certain. When $A$ is a classical event, the above inequality strengthens into $N(A) > 0 \Rightarrow \Pi(A) = 1$. Thus, when $A$ is a classical event and when $\pi$ is normalized, one can distinguish between the five following epistemic states of knowledge:

1. $N(A) = 1 = \Pi(A)$ : Given that "$V$ is $E$", $A$ is *certain* ("$V$ is $A$" is true in all non-impossible situations, knowing that "$V$ is $E$") ;

2. $0 < N(A) < 1$ and $\Pi(A) = 1$ : Given that "$V$ is $E$", $A$ is normally true ("$V$ is $A$" is true in all the most plausible situations)

3. $N(A) = 0$ and $\Pi(A) = 1$ : this is a state of *complete ignorance* about $A$, since both $A$ and $\overline{A}$ are completely possible ($\Pi(A) = 1 = \Pi(\overline{A})$). Note that we should have here $A \neq U$, since $N(U) = 1$ and $\Pi(\emptyset) = 0$.

4. $N(A) = 0$ and $0 < \Pi(A) < 1$ : Given that "$V$ is $E$", $A$ is normally false (all situations $u$ where "$V$ is $A$" is true are somewhat impossible, i.e. $\pi(u) < 1$)

5. $N(A) = 0$ and $\Pi(A) = 0$ : Given that "$V$ is $E$", $A$ is *certainly false* ("$\forall u \in A, \pi(u) = 0$, and then "$V$ is $A$" is false and equivalently, "$V$ is $\overline{A}$" is true in all non-impossible situations $u$, that is in all situations somewhat compatible with "$V$ is $E$")

Besides, complete ignorance about any non-trivial event $A \neq \emptyset$ is represented by $\forall u \in U, \pi(u) = 1$.

## Guaranteed possibility and potential certainty measures

Apart from $\Pi$ and $N$, two other set-functions can be defined using "max" or "min", namely (see e.g.[10]):

- a measure of "guaranteed possibility"

$$\Delta(A) = \min_{u \in A} \pi(u),$$

  which estimates to what extent all elements in $A$ are possible. Clearly $\Delta$ is a stronger measure than $\Pi$, i.e., $\Delta \leq \Pi$, since $\Pi$ only estimates the existence of at least one value in $A$ compatible with the available knowledge, while the evaluation provided by $\Delta$ concerns all the values in $A$. Note also that $\Delta(A)$ and $N(A)$ are unrelated.

- a dual measure of "potential certainty"

$$\nabla(A) = 1 - \Delta(\overline{A}) = 1 - \min_{u \notin A}(\pi(u))$$

  which estimates to what extent there exists at least one value in the complement of $A$ that has a low degree of possibility; this is a necessary condition for having "$u \in A$" somewhat certain (but in general far from being sufficient, except if $\overline{A}$ has only one element).

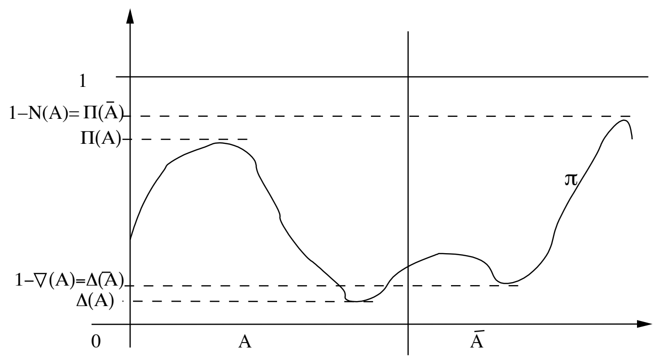


Figure 5. The four values associated to an event $A$ w.r.t. a possibility distribution $\pi$.

The figure 5 shows the four values $\Pi$, $N$, $\Delta$ and $\nabla$ associated to an event $A$ given a possibility distribution $\pi$. Obviously we have $N \leq \nabla$. Moreover, if there exists at least one value in $U$ which is impossible w.r.t. $\pi$, (i.e., $\exists u \in U, \pi(u) = 0$) the constraint $\min(\Delta(A), \Delta(\overline{A})) = 0$ holds and then $\Delta(A) > 0$ entails $\nabla(A) = 1$ (at the technical level, it is always possible to add an element $u_0$ to $U$ such that $\pi(u_0) = 0$).

$\Delta$ and $\nabla$ are monotonically decreasing set functions (in the wide sense) with respect to set inclusion, it contrasts with $\Pi$ and $N$ which are monotonically increasing. Indeed, they satisfy the following equalities:

$$\Delta(A \cup B) = \min(\Delta(A), \Delta(B)) \text{ and } \nabla(A \cap B) = \max(\nabla(A), \nabla(B)).$$

Like for the possibility and necessity measures, we only have the inequalities:

$$\Delta(A \cap B) \geq \max(\Delta(A), \Delta(B)) \text{ and } \nabla(A \cup B) \leq \min(\nabla(A), \nabla(B)).$$

The characteristic property of $\Delta$ may be interpreted as "$A$ or $B$" is allowed, permitted if and only if $A$ is permitted and $B$ is permitted. So $\Delta$ functions may model explicit permission in a deontic framework.

Note that the four quantities $\Pi(A)$, $N(A)$, $\Delta(A)$ and $\nabla(A)$ are only weakly related since they are only constrained by

$$\begin{array}{lll} \max(\Pi(A), 1 - N(A)) & = \max_{u \in U} \pi(u) & (= 1 \text{ if } \pi \text{ is normalized}) \\ \min(\Delta(A), 1 - \nabla(A)) & = \min_{u \in U} \pi(u) & (= 0 \text{ if } 1 - \pi \text{ is normalized}) \end{array}$$

together with the duality relations between $N$ and $\Pi$, and between $\Delta$ and $\nabla$. Moreover we have the following inequality (proved in [9]) when both $\pi$ and $1 - \pi$ are normalized:

**Proposition 7.1.** If $\pi$ and $1 - \pi$ are normalized then

$$\max(N(A), \Delta(A)) \leq \min(\Pi(A), \nabla(A)).$$

Note that this inequality agrees with the intuition: i) the higher the certainty of $A$, the higher the possibility of $A$, i.e., the consistency of $A$ with what is known, and the higher its potential certainty; ii) the higher the feasibility of all the values in $A$ in the sense of $\Delta$, the higher the possibility of $A$, and the higher the potential certainty of $A$.

### Conditioning in possibility theory

In the qualitative possibility theory setting, conditioning is defined from the following relation

$$\Pi(A \cap B) = \min(\Pi(B|A), \Pi(A))$$

in order to express in a qualitative way that having "A and B true" is as possible as having "$A$ true" possible and having "$B$ true" possible in the context where "$A$ is true". When $\Pi(A) > 0$, $\Pi(B|A)$ is defined from the above equation by looking for its greatest solution and by applying the minimal specificity principle. Namely, since $\Pi(A \cap B) \leq \Pi(A)$ always holds by monotony of the set function $\Pi$, we have

$$\Pi(B|A) = \begin{cases} 1 & \text{if} \quad \Pi(A \cap B) = \Pi(A) > 0 \\ \Pi(A \cap B) & \text{if} \quad \Pi(A \cap B) < \Pi(A) \end{cases}$$

Then, the conditional necessity is defined by duality

$$N(B|A) = 1 - \Pi(\overline{B}|A)$$

Then it can be checked that

$$N(B|A) > 0 \quad \Leftrightarrow \quad \Pi(A \cap B) > \Pi(A \cap \overline{B}),$$

which expresses that $B$ is somewhat certain when knowing that $A$ is true if and only if having $A \cap B$ true is strictly more possible than having $A \cap \overline{B}$ true. This is the basis for representing default rules of the form "if $A$ is true then generally $B$ is true" by the above inequality [3].

Similarly,

$$\Delta(A \cap B) = \max(\Delta(B|A), \Delta(A))$$

and applying the *maximal* specificity principle (it works in a reverse way with respect to the $\Pi$-situation), we obtain if $\Delta(A) < 1$

$$\Delta(B|A) = \begin{cases} 0 & \text{if} \quad \Delta(A \cap B) = \Delta(A) < 1 \\ \Delta(A \cap B) & \text{if} \quad \Delta(A \cap B) > \Delta(A) \end{cases}$$

It can be checked that

$$\Delta(B|A) > 0 \Leftrightarrow \Delta(A \cap B) > \Delta(A \cap \overline{B})$$

## References

[1] Barbut, M., Montjardet, B.: *Ordre et Classification: Algèbre et Combinatoire*, Hachette, 1970.

[2] Belohlavek, R.: *Fuzzy Relational Systems: Foundations and Principles*, vol. 20, Kluwer Academic/Plenum Press, New York, 2002.

[3] Benferhat, S., Dubois, D., Prade, H.: Representing default rules in possibilistic logic, *Proc. of the 3 $^{rd}$ Conf. on Knowledge Representation*, 1992.

[4] Birkhoff, G.: *Lattice Theory*, American Mathematical Society Colloquium Publications vol. XXV, 1940.

[5] Burusco, A., Fuentes-Gonzlez, R.: Construction of the L-fuzzy concept lattice, *Fuzzy Sets and Systems*, **97**(1), July 1998, 109–114.

[6] Dubois, D., Fariñas Del Cerro, L., Herzig, A., Prade, H.: A roadmap of qualitative independence , in: *Fuzzy Sets, Logics and Reasoning about Knowledge* (D. Dubois, H. Prade, E. Klement, Eds.), vol. 15 of *Applied Logic series*, Kluwer Academic Publishers, Dordrecht, Pays-Bas, 1999, 325–350.

[7] Dubois, D., Prade, H.: *Possibility Theory: An Approach to Computerized Processing of Uncertainty*, Plenum Press, New York, 1988, With the collaboration of H. Farreny, R. Martin-Clouaire, C. Testemale.

[8] Dubois, D., Prade, H.: Rough fuzzy sets and fuzzy rough sets, *Int. Journal of General Systems*, **17**(2-3), 1990, 191–209.

[9] Dubois, D., Prade, H.: Possiblity theory as a basis for preference propagation in automated reasonning, *Proc. of the 1st IEEE Inter. Conf. on Fuzzy Systems (FUZZ-IEEE'92)*, San Diego, CA, March 1992.

[10] Dubois, D., Prade, H.: Possibility theory:qualitative and quantitative aspects, in: *Quantified Representation of Uncertainty and Imprecision*, vol. 1 of *Handbook of Defeasible Reasoning and Uncertainty Management Systems*, Kluwer Academic Publishers, The Netherlands, 1998, 169–226.

[11] Düntsch, I., Gediga, G.: Approximation operators in qualitative data analysis, in: *Theory and application of relational structures as knowledge instruments*, Springer, Heidelberg, 2003, 216–233.

[12] Düntsch, I., Orlowska, E.: Mixing modal and sufficiency operators, *Bulletin of the Section of Logic, Polish Academy of Sciences*, **28**(2), 1999, 99–106.

[13] Düntsch, I., Orlowska, E.: Boolean algebras arising from information systems, *Annals of Pure and Applied Logic*, **127**(Issues 1-3), June 2004, 77–98.

[14] Ganter, B.: *Two basic algorithms in concept analysis*, Technical Report Preprint 831, Technische Hochschule Darmstadt, 1984, 28 p.

[15] Ganter, B., Wille, R.: *Formal Concept Analysis, Mathematical Foundations*, Springer-Verlag, 1999.

[16] Gediga, G., Düntsch, I.: Modal-style operators in qualitative data analysis, *Proceedings of the 2002 IEEE International Conference on Data Mining*, December 9 - 12 2002.

[17] McCarthy, J., Hayes, P.: Some philosophical problems from the standpoint of artificial intelligence, in: *Machine Intelligence*, vol. 4, 1969, 463–502.

[18] Nahmias, S.: Fuzzy variables, *Fuzzy Sets and Systems*, **1**(2), 1978, 97–110.

[19] Ore, O.: Galois Connexions, *Transactions of the American Mathematical Society*, **55**, 1944, 493–513.

[20] Pawlak, Z.: Rough sets, *Internat. J. Computer and Information Sciences*, **11**, 1982, 341–356.

[21] Pawlak, Z.: *Rough Sets. Theoretical Aspects of. Reasoning about Data*, Kluwer Acad. Publ., Dordrecht, 1991.

[22] Wille, R.: Restructuring lattice theory: an approach based on hierarchies of concepts, in: *Ordered Sets* (I. Rival, Ed.), D. Reidel, Dordrecht-Boston, 1982, 445–470.

[23] Yao, Y.: A comparative study of formal concept analysis and rough set theory in data analysis, *Rough Sets and Current Trends in Computing, 4th International Conference, RSCTC 2004*, Uppsala, Sweden, June 1-5 2004.

[24] Yao, Y.: Concept lattices in Rough set theory, *Proc. of Fuzzy Sets in the Heart of Canadian Rockies conference (NAFIPS-04)*, June 27-30 2004.

[25] Zadeh, L.: Fuzzy sets as a basis for a theory of possibility, *Fuzzy Sets and Systems*, **1**, 1978, 3–28.