An independence concept under plausibility function

Marcello Mastroleo
Dip. Matematica e Informatica
Università di Perugia
mastroleo@dipmat.unipg.it

Barbara Vantaggi
Dip. Metodi e Modelli Matematici
Università "La Sapienza" Roma
vantaggi@dmmm.uniroma1.it

Abstract

Starting from considering different definitions of conditioning for decomposable measures, in particular for totaly monotone measures (belief functions) and totaly alternating measures (plausibility functions), we provide a concept of independence which covers some natural properties. In particular, we characterize the proposed independence for plausibility functions and we check some relevant properties. Relationships with other notions studied in literature are shown.

Keywords. Totaly monotone measures, Plausibility, Conditioning, Independence.

1 Introduction

The subtle notion of conditioning is controversial in several contexts, for example for non-additive measures and more specifically for plausibility and belief functions (which are totally alternating and monotone, respectively). We consider a general axiomatic definition of conditional measures proposed in [7]: the conditional measure is *directly* defined as a function on a set of conditional events which satisfies a suitable set of axioms. In this framework conditional measures are seen as a primitive notion, analogously to conditional probability according to de Finetti approach [16]. Among the conditional measures we deal with conditional plausibility and belief functions.

The theory of belief functions, also known as Dempster-Shafer theory [18] and theory of evidence, aims to model degree of belief. It can be regarded as a generalization of the probability approach and many interpretation have been proposed: a belief function can be seen as a particular lower probability or it can be derived from probability where a probability space is mapped by a one-to-many mapping on another space.

In [10] it is shown a sort of converse property of the fact that a belief function is a specific lower probability: a lower probability obtained as extension of a suitable coherent conditional probability is a belief function.

Starting from this general framework some other well-known definitions arise naturally (see also [6]). A comparison of these different conditioning operators has been carried out in [12] from another point of view by looking to a comparative setting and, more precisely, by studying local representability of ordinal relations defined on a finite algebra.

In particular, we refer to a well-known definition (see [19, 26]) of conditional belief, which can be obtained from the above one as particular case and, for any given conditioning event H, it can be seen as the dual function of a conditional plausibility.

Actually, we refer to a generalization of the definition provided in [26] that allows to deal also with events of zero plausibility. Then, the problem of dealing with "partial assessments" on (not necessarily structured) domains, containing only elements of interest, is faced. In any real situation the events of interest, and those in which the field expert or the decision maker has information, give rise usually to an arbitrary set. For this reason we need a notion of consistency, which allows to check whether a partial assessment is the restriction of a conditional belief function (or a conditional plausibility) [6]. A characterization of both consistent conditional plausibility and conditional belief, in terms of a suitable class of plausibility functions, is carried out: a conditional belief/plausibility is not always singled out by a unique unconditional measure.

In such framework, we study an important concept for uncertainty reasoning, which is independence. In probabilistic theory this condition has been deeply studied (see e.g. [15, 9, 32]); moreover such notion has been studied also in other non-probabilistic frameworks [1, 4, 11, 30, 31, 35] and in particular in upper and lower probabilities theory (see, for example, [5, 8, 13, 34]).

However, the concept of independence has not been widely treated in belief theory (see [2, 3, 27, 29]). In addition to the theoretical reasons for the study of independence, there are also practical interest: many computational tasks can be simplified by using independence notion.

In this paper we propose a definition of independence for conditional plausibility, which can be reformulated by means of duality property also for belief functions. This notion covers some natural properties also in the case of events with degree equal to 0 or 1. In particular, we show that such independence notion implies logical independence. This is an intuitive implication: in fact *if an event is "logically" related to another one, the two events must be dependent* under *any* uncertainty measure. Handling logical constraints is interesting also from a practical point of view, since in many real applications (e.g. in finance, economics, medicine) variables are suitably linked (see e.g. [20, 21]). Then, we get this natural implication, which does not need to be required explicitly as in [2]. Actually, our definition of independence is inspirited to that one given for coherent conditional probability in [9, 32] and that for conditional possibilities in [11, 24].

We give a numerical characterization of the proposed definition of independence that helps to compare our definition with other ones given in literature [2, 3, 27, 29].

In Section 2, we introduce conditional plausibility and in Section 2.1 (through duality) conditional belief; we briefly deal with a consistency notion for partial assessments.

In Section 3 we provide an independence notion firstly for plausibility and then for belief function. We study its main properties by comparing it also with other notions introduced in literature.

2 Conditional plausibility and belief functions

Usually in literature conditional measures are presented as a derived notion of unconditional ones, but this is a restrictive view of conditioning. It is instead essential to adopt a general definition of generalized $(\oplus, \odot)$-decomposable conditional (uncertainty) measure (introduced in [10]). The peculiarity of this approach consists in the fact that conditional measures are directly defined on a suitable set of conditional events.

Moreover, by specifying the two operations $\oplus, \odot$ we obtain some particular conditional uncertainty measures. In particular, by taking the usual sum and product, respectively, we get a conditional probability (in the sense of de Finetti [16]), while for $\oplus = \max$ and $\odot = \min$ we obtain conditional possibility [4, 23]. In [10] it is shown that for specific operations conditional belief functions can be seen as particular generalized decomposable measures.

In order to revisit the belief functions and their connections with the so-called "imprecise" probabilities and with extensions of coherent conditional probabilities, we recall firstly some basic notions and then some results given in [10]. An assessment p on a set of conditional events $\mathcal{C}$ is a coherent conditional probability iff there exists a conditional probability P on the product of an algebra $\mathcal{E}$ and an additive set (closed under finite unions) $\mathcal{H} \subseteq \mathcal{E} \setminus \{\emptyset\}$ such that $\mathcal{C} \subseteq \mathcal{E} \times \mathcal{H}$ and the restriction of P on $\mathcal{C}$ coincides with p (i.e. $P(E|H) = p(E|H)$ for any $E|H \in \mathcal{C}$).

Given an *arbitrary* set $\mathcal{C}$ of conditional events, a *coherent lower conditional probability* on $\mathcal{C}$ is a nonnegative function $\underline{P}$ such that there exists a non-empty *dominating family* $\mathcal{P} = \{P(\cdot|\cdot)\}$ of *coherent* conditional probabilities on $\mathcal{C}$ whose lower envelope is $\underline{P}$, that is, for every $E|H \in \mathcal{C}$,

$$\underline{P}(E|H) = \inf_{\mathcal{P}} P(E|H)\,.$$

In particular, by taking $\mathcal{C}$ as a set of unconditional events, we get a coherent lower probability.

It is well known that a belief function (totaly monotone measure) is a lower probability; the following result proved in [10] shows the converse property: a lower probability obtained as extension of a suitable coherent conditional probability is a belief function.

Theorem 1 *Let $\mathcal{D} = \{H_1, \ldots, H_n\}$ be a finite set of pairwise incompatible events. Denoting by $\mathcal{K}$ the additive set spanned by them, and given an algebra $\mathcal{A} \supset \mathcal{K}$, put $\mathcal{C} = \mathcal{A} \times \mathcal{K}$. If $P(\cdot)$ is a coherent probability on $\mathcal{D}$, let $\mathcal{P}$ be the class of coherent conditional probabilities $P(\cdot|\cdot)$ extending $P(\cdot)$ on $\mathcal{C}$. Consider, for $E|K \in \mathcal{C}$, the lower probability*

$$\underline{P}(E|K) = \inf_{\mathcal{P}} P(E|K)\,; \tag{1}$$

then for any $K \in \mathcal{K}$ the function $\underline{P}(\cdot\,|K)$ is a belief function on $\mathcal{A}$.

The involved set $\mathcal{D}$ is not a consequence of some particular circumstances, but it is always possible to find it, as shown by the following theorem [10] (Section 3.1.):

Theorem 2 *Let $\mathcal{A}$ be a finite algebra and Bel be a belief function on $\mathcal{A}$. Then, there exists a partition*

$\mathcal{D} = \{H_1, \dots, H_n\}$ *of* Ω *and a (coherent) probability on* $\mathcal{D}$ *such that the lower envelope of the class of coherent conditional probabilities* $P(\cdot|\cdot)$ *extending* $P(\cdot)$ *on* $\mathcal{C} = \mathcal{A} \times \mathcal{K}$ *(*$\mathcal{K}$ *is the additive set generated by* $\mathcal{D}$*) coincides with* Bel *on* $\mathcal{A}$.

The following axioms are naturally derived (see [6]):

Definition 1 *Let* $\mathcal{E}$ *be an algebra and* $\mathcal{H} \subseteq \mathcal{E} \setminus \{\emptyset\}$ *an additive set. A function* Pl *defined on* $\mathcal{C} = \mathcal{E} \times \mathcal{H}$ *is a* conditional plausibility *if it satisfies the following conditions*

i) $Pl(E|H) = Pl(E \wedge H|H)$;

ii) $Pl(\cdot|H)$ *is a plausibility function* $\forall H \in \mathcal{H}$;

iii) For every $E \in \mathcal{E}$ *and* $H, K \in \mathcal{H}$

$$Pl(E \wedge H|K) = Pl(E|H \wedge K) \cdot Pl(H|K).$$

Moreover, given a conditional plausibility, a conditional belief *function* $Bel(\cdot|\cdot)$ *is defined by duality as follows: for every event* $E|H \in \mathcal{C}$

$$Bel(E|H) = 1 - Pl(E^c|H).$$

It is possible to see that the above axiomatization extends the Dempster's rule, i.e.

$$Bel(F|H) = 1 - \frac{Pl(F^c \wedge H)}{Pl(H)},$$

for all H such that $Pl(H) > 0$ (from condition *iii)*). When all the conditioning events have positive plausibility, i.e. $\Omega \in \mathcal{H}$ and $Pl(H|\Omega) > 0$ for any $H \in \mathcal{H}$, the above notions of conditional plausibility and conditional belief coincide with that given in [19, 26]. In fact, if $Pl(H) > 0$ it follows $Bel(F|H) = \frac{Pl(H) - Pl(F^c \wedge H)}{Pl(H)} = \frac{Bel(F \vee H^c) - Bel(H^c)}{Pl(H)}$.

2.1 Coherent conditional belief

By regarding a conditional plausibility function as a $(\oplus, \odot)$-decomposable measure, it is possible to study the structure underlying the conditional measure and to build an algorithm to check the consistency (with the model of reference) of a partial assessment.

In the following we denote by $\mathcal{F} = \{E_1|F_1, E_2|F_2, \dots, E_m|F_m\}$ an arbitrary finite set of conditional events, by $\mathcal{E}$ the algebra generated by $\{E_1, F_1, \dots, E_m, F_m\}$ and by $\mathcal{K}$ the additive set generated by the set of the conditioning events $\{F_1, \dots, F_m\}$.

Definition 2 *A function* $f(\cdot|\cdot)$ *on an arbitrary finite set* $\mathcal{F}$ *is a* coherent *conditional belief (plausibility) if there exists* $\mathcal{C} \supset \mathcal{F}$*, with* $\mathcal{C} = \mathcal{E} \times \mathcal{K}$ *such that* $f(\cdot|\cdot)$ *can be extended from* $\mathcal{F}$ *to* $\mathcal{C}$ *as a* conditional belief (conditional plausibility).

The following theorem [6] characterizes (coherent) conditional belief functions in terms of a class of plausibilities $\{Pl_1, ..., Pl_m\}$.

Theorem 3 *Let* $\mathcal{F} = \{E_1|F_1, E_2|F_2, \dots, E_m|F_m\}$ *be an arbitrary finite set of conditional events and denote by* $\mathcal{E} = \{H_1, H_2, ..., H_n\}$ *the algebra generated by* $\{E_1, \dots, E_m, F_1, \dots, F_m\}$ *and* $H_0^0 = \vee_{j=1}^m F_j$*. For a real function* Bel *on* $\mathcal{F}$ *the following statements are equivalent:*

(a) $Bel : \mathcal{F} \to [0, 1]$ *is a coherent conditional belief assessment;*

(b) there exists (at least) a class $\mathcal{L} = \{Pl_\alpha\}$ *of plausibility functions such that* $Pl_\alpha(H_0^\alpha) = 1$ *and* $H_0^\alpha \subset H_0^\beta$ *for all* $\beta < \alpha$*, where* H_0^α *is the greatest element of* $\mathcal{K}$ *for which* $Pl_{(\alpha-1)}(H_0^\alpha) = 0$.

Moreover, for every $E_i|F_i$*, there exists an index* α *such that* $Pl_\beta(F_i) = 0$ *for all* $\alpha > \beta$*,* $Pl_\alpha(F_i) > 0$ *and*

$$Bel(E_i|F_i) = 1 - \frac{Pl_\alpha(E_i^c F_i)}{Pl_\alpha(F_i)}, \qquad (2)$$

(c) all the following systems (S^α)*, with* $\alpha = 0, 1, 2, ..., k \leq n$*, admit a solution* $\mathbf{X}^\alpha = x_k^\alpha = m_\alpha(H_k)$*:*

$$(S^\alpha) = \begin{cases} \sum\limits_{H_k F_i \neq \emptyset} x_k^\alpha \cdot [1 - Bel(E_i|F_i)] = \sum\limits_{H_k E_i^c F_i \neq \emptyset} x_k^\alpha, \ \forall F_i \subseteq H_0^\alpha \\ \sum\limits_{H_k \in H_0^\alpha} x_k^\alpha = 1 \\ x_k^\alpha \geq 0, \qquad \forall H_k \subseteq H_0^\alpha \end{cases}$$

where H_0^α *is the greatest element of* $\mathcal{K}$ *such that* $\sum\limits_{H_i H_0^\alpha \neq \emptyset} m_{(\alpha-1)}(H_i) = 0$.

The above characterization result holds for coherent conditional belief functions as well as for coherent conditional plausibility. In particular condition *(c)* stresses that this measure can be written in terms of a suitable class of basic assignments, instead of just one as in the classical case where all the conditioning events have positive plausibility.

Note that every class $\mathcal{L}$ (condition (b) of Theorem 3) is said to be agreeing with both the conditional belief Bel and its dual conditional plausibility Pl. Whenever there are events in $\mathcal{K}$ with zero plausibility the class of unconditional plausibilities is formed by more than one element and we can say that Pl_1 gives a refinement of those events judged with zero plausibility under Pl_0.

The following example shows the construction of the class $\mathcal{L}$ characterizing (in the sense of the above result) a conditional belief.

Example 1 *Let* $\{C_1, ..., C_5\}$ *be a partition of* Ω, $\mathcal{E}$ *the corresponding algebra and* $\mathcal{K} = \{C_1 \vee C_5, C_2 \vee C_3 \vee C_4, C_1 \vee C_2 \vee C_5, \Omega\}$.

Consider the following function f *defined as follows on* $\mathcal{E} \times \mathcal{H}$:
for $K \in \{\Omega, C_2 \vee C_3 \vee C_4\}$ *and* $H \subseteq C_1 \vee C_5$
$f(C_i|K) = f(H|K) = f(H \vee C_i|K) = 0$ *for* $i = 3, 4$
$f(C_2|K) = f(C_2 \vee H|K) = f(C_2 \vee C_4|K) = f(C_2 \vee C_4 \vee H|K) = 0.5$,
$f(C_3 \vee C_4|K) = f(C_3 \vee C_4 \vee H|K) = 0.2$,
$f(C_2 \vee C_3|K) = f(C_2 \vee C_3 \vee H|K) = 0.8$
$f(C_2 \vee C_3 \vee C_4|K) = f(C_2 \vee C_3 \vee C_4 \vee H|K) = 1$;

moreover (for $i = 1, 5$*)*
$f(C_i|C_1 \vee C_2 \vee C_5) = f(C_1 \vee C_5|C_1 \vee C_2 \vee C_5) = 0$,
$f(C_2|C_1 \vee C_2 \vee C_5) = f(C_2 \vee C_i|C_1 \vee C_2 \vee C_5) = f(C_1 \vee C_2 \vee C_5|C_1 \vee C_2 \vee C_5) = 1$;

and $f(C_1|C_1 \vee C_5) = 0.2$, $f(C_5|C_1 \vee C_5) = 0.3$, $f(C_1 \vee C_5|C_1 \vee C_5) = 1$.

We can prove that the above function is a conditional belief since there exists a suitable class $\mathcal{L} = \{Pl_0, Pl_1\}$ *of plausibilities such that, for any* $E|F \in \mathcal{A} \times \mathcal{K}$, *one has* $f(E|F) = 1 - \frac{Pl_\alpha(E^c \wedge F)}{Pl_\alpha(F)}$. *The function* Pl_0 *is defined on* $\mathcal{A}$ *as follows: for any* $H \subseteq C_1 \vee C_5$ $Pl_0(H) = 0$, $Pl_0(C_2) = Pl_0(C_2 \vee H) = 0.8$, $Pl_0(C_4) = Pl_0(C_4 \vee H) = 0.2, Pl_0(C_3) = Pl_0(C_3 \vee H) = Pl_0(C_3 \vee C_4) = Pl_0(C_3 \vee C_4 \vee H) = 0.5$,
$Pl_0(C_2 \vee C_3) = Pl_0(C_2 \vee C_3 \vee H) = Pl_0(C_2 \vee C_4) = Pl_0(C_2 \vee C_4 \vee H) = Pl_0(C_2 \vee C_3 \vee C_4) = Pl_0(C_2 \vee C_3 \vee C_4 \vee H) = 1$.
Note that Pl_0 *is associated to the following basic assignment* $m(C_2) = 0.5, m(C_2 \vee C_3) = 0.3, m(C_3 \vee C_4) = 0.2$ *and it is zero otherwise.*
Then, $H_1^0 = C_1 \vee C_5$, *and* Pl_1 *is defined as follows* $Pl_1(C_1) = 0.7, Pl_1(C_5) = 0.8, Pl_1(C_1 \vee C_5) = 1$.

Results similar to the above one, characterizing conditional possibility and necessity in terms of a class of unconditional possibilities, have been given in [4, 11, 24], and for conditional probability see e.g. [9].

2.2 Zero–layers

The characterization of conditional plausibility (and conditional belief function) in terms of a suitable class of plausibilities gives rise to the following notion of zero-layers.

Definition 3 *Let* Pl *be a coherent conditional plausibility on* $\mathcal{F}$, *and* $\mathcal{L}$ *a class agreeing with* Pl, *then, for every event* $H \in \mathcal{E}$, *the* zero–layer *of* H *(denoted as* $\circ(H)$*) related to* $\mathcal{L}$ *is defined as the minimum number* α *such that* $Pl_\alpha(H) > 0$.
Moreover, define $\circ(\emptyset) = +\infty$.

Zero-layers single-out a partition of the algebra, in particular it follows that the zero-layer of any event E with positive plausibility is zero. Then, if the class $\mathcal{L}$ contains only an everywhere positive plausibility Pl_o, there is only one (trivial) zero-layer.

Remark 1 *It is immediate to prove that the zero-layers, related to* $\mathcal{L}$, *satisfy the following formal proprieties*

$$\circ(A \vee B) = \min\{\circ(A), \circ(B)\},$$
$$\circ(A \wedge B) \geq \max\{\circ(A), \circ(B)\}.$$

Note that zero-layers (which are obviously significant for events of zero plausibility) are a tool to detect "how much" a null event is ... null. In fact, if $\circ(A) > \circ(B)$ (that is, roughly speaking, the plausibility of A is a "stronger" zero than the plausibility of B), then by Theorem 3 (b) $Pl(A|A \vee B) = 0$ and so $Pl(B|A \vee B) = 1$. On the other hand $\circ(A) = \circ(B)$ iff $Pl(A|A \vee B)Pl(B|A \vee B) > 0$; this formula recalls the probabilistic notion of *commensurable* given by de Finetti in [17].

Definition 4 *Let* Pl *be a coherent conditional plausibility on* $\mathcal{F}$, *and* $\mathcal{L}$ *a class agreeing with* Pl, *then, for every event* $E|H \in \mathcal{E} \times \mathcal{K}$, *the* zero–layer *of* $E|H$ *(denoted as* $\circ(E|H)$*) related to* $\mathcal{L}$ *is defined as the (positive) number*

$$\circ(E|H) = \circ(E \wedge H) - \circ(H).$$

Since $\circ(\emptyset) = \infty$ it results $\circ(E|H) = \infty$ iff $E \wedge H = \emptyset$.

Remark 2 *More precisely,* $Pl(A|B) > 0$ *if and only if* $\circ(A|B) = 0$ *(i.e.* $\circ(A \wedge B) = \circ(B)$*).*

Moreover, from the properties of conditional plausibilities, for any conditioning event H, there is at least an atom $C \subseteq H$ such that $\circ(C|H) = 0$.

Example 1 (continued) Let us consider again the conditional plausibility in Example 1, which admits a unique agreeing class and note that $\circ(C_1 \vee C_5) = 1$ and $\circ(C_1|C_1 \vee C_5) = \circ(C_5|C_1 \vee C_5) = 0$.

The above properties recall those related to the notion of zero-layer [9] arising in de Finetti conditional probability framework and they satisfy the same properties of k-functions of Spohn [30], so suggest relevant connections with the results shown in [11, 22, 24].

3 Independence

The background is now ready to introduce a definition of independence for coherent conditional plausibilities (i.e. the measure can be assessed on arbitrary set of conditional events without requiring any algebraic structure).

Definition 5 *Given a coherent conditional plausibility Pl on a set of conditional events $\mathcal{F}$ containing $\mathcal{D} = \{A^*|B^*, A^*\}$ - where A^* (analogously B^*) stands for either A or A^c-, A is independent of B under Pl (in symbol $A \perp\!\!\!\perp B[Pl]$), if both the following condition holds:*

(a) $Pl(A|B) = Pl(A|B^c) = Pl(A)$
$Pl(A^c|B) = Pl(A^c|B^c) = Pl(A^c)$,

(aa) there exists an agreeing class $\mathcal{L} = \{Pl_\alpha\}$ for the restriction of Pl to $\mathcal{D}$ such that

$$\circ(A|B) = \circ(A|B^c) \text{ and } \circ(A^c|B) = \circ(A^c|B^c).$$

Remark 3 *Definition 5 requires for the statement "A independent of B under [Pl]" that $B \neq \Omega$ and $B \neq \emptyset$ (since conditioning events cannot be impossible).*

This syntactical constraint has also a semantical counterpart: Ω and $\emptyset$ correspond to a situation of complete information (since the former is always true and the latter always false), and so it does not make sense to ask whether they could influence the plausibility of another event.

Conversely, by definition it follows that, under any coherent conditional plausibility, the events Ω and $\emptyset$ are independent of every possible (i.e. different from Ω and $\emptyset$) event B. In fact, condition (i) holds and for any agreeing class $\circ(\Omega|B) = \circ(\Omega|B^c) = 0$ and $\circ(\emptyset|B) = \circ(\emptyset|B^c) = +\infty$.

This conclusion is natural, since the plausibility (1 and 0, respectively) of Ω and $\emptyset$ cannot be changed by assuming the occurrence of any other possible event B.

In condition *(a)* of Definition 5 we require equalities that could seem very strong at the first light, this is due to remove situations such as those arising in the following examples:

Example **2** *Let consider a basic assignment on the algebra generated by two possible events A and B, with focal elements*

$$m(A \wedge B) = m(A \wedge B^c) = m(A^c \wedge B) =$$

$$m(A^c \wedge B^c) = m(A \vee B) = \frac{1}{5}$$

(i.e. on all the other events of the algebra $m(\cdot)$ is equal to zero). This basic probability assignment implies $Pl(A) = Pl(A^c) = Pl(B) = Pl(B^c) = \frac{3}{5}$, $Pl(A \wedge B) = Pl(A^c \wedge B) = Pl(A \wedge B^c) = \frac{2}{5}$ but $Pl(A^c \wedge B^c) = \frac{1}{5}$. By applying the conditioning rule (Definition 1) it follows that $Pl(A|B) = Pl(A|B^c) = \frac{2}{3} \neq Pl(A)$. Moreover, $Pl(A^c|B) = \frac{2}{3} \neq \frac{1}{3} = Pl(A^c|B^c)$.

The above example shows that $Pl(A|B) = Pl(A|B^c)$ does not imply neither $Pl(A|B) = Pl(A)$ nor $Pl(A^c|B) = Pl(A^c|B^c)$, furthermore from the next example it arises the necessity of requiring all the equalities in condition *(a)*.

Example **3** *Consider the following basic assignment*

$$m(A \wedge B) = m(A \wedge B^c) = m(A^c \wedge B) =$$

$$m(A^c \wedge B^c) = m(\Omega) = \frac{1}{5}.$$

Then, $Pl(A^* \wedge B^*) = \frac{2}{5}, Pl(A^*) = Pl(B^*) = \frac{3}{5}$ and $Pl(A^*|B^*) = \frac{2}{3}$. This implies that

$$Pl(A|B) = Pl(A|B^c)$$

and

$$Pl(A^c|B) = Pl(A^c|B^c),$$

but $Pl(A|B) \neq Pl(A)$.

When both $Pl(A)$ and $Pl(A^c)$ are greater than zero condition *(a)* of Definition 5 assures that $A \perp\!\!\!\perp B[Pl]$, in fact in this case all the zero-layers in condition *(aa)* are equal to 0 and so condition *(aa)* is trivially satisfied.

If condition *(a)* holds and $Pl(A) = 0$ [$Pl(A^c) = 0$], then the second [first] equality under *(aa)* is trivially satisfied, so that the statement $A \perp\!\!\!\perp B[Pl]$ is ruled by the first [second] one. In other words equality *(a)* is not enough to assure independence in this situation: *it needs to be reinforced by the requirement that also their zero-layers must be equal.*

We finally note that the statement $A \perp\!\!\!\perp B\,[Pl]$ depends only on the restriction of the assessment Pl on $\mathcal{D}$, hence the statement is not effected by the values of the assessment Pl on $\mathcal{F} \setminus \mathcal{D}$ (actually the influence, e.g. of $Pl(B|A)$ is related to condition *(aa)*, as it will be clear from the next result). Since *(aa)* depends on a class agreeing with the coherent conditional plausibility, and since this class is in general not unique, it is necessary to prove that independence is well-defined by Definition 5, that means that is invariant with respect to the choice of any agreeing class.

Theorem 4 *Given two events A and B such that $B \neq \emptyset, \Omega$ and a coherent conditional plausibility Pl defined on $\mathcal{F}$, containing $\mathcal{D} = \{A^*|B^*, A^*\}$, such that*

$$Pl(A|B) = Pl(A|B^c) = Pl(A)$$
$$Pl(A^c|B) = Pl(A^c|B^c) = Pl(A^c).$$

If there exists a class agreeing with $Pl|_{\mathcal{D}}$ such that

$$\circ(A|B) = \circ(A|B^c) \text{ and } \circ(A^c|B) = \circ(A^c|B^c),$$

then this holds for any other class agreeing with $Pl|_{\mathcal{D}}$.

Proof: This theorem can be decomposed in three main cases:

1. $Pl(A) \cdot Pl(A^c) > 0$,
2. $Pl(A) = 0$,
3. $Pl(A^c) = 0$.

1. If $Pl(A) \cdot Pl(A^c) > 0$ the theorem is true since $\circ(A^*|B^*) = 0$ for all agreeing class.

2. If $Pl(A) = 0$ then $Pl(A^c) = 1$ and the only masses which can be greater than zero (i.e. the focal elements) are $m(A^c \wedge B)$, $m(A^c \wedge B^c)$, $m(A^c)$. If an agreeing class is such that $Pl(B) \cdot Pl(B^c) > 0$ (i.e. $m(A^c) > 0$ or $m(A^c \wedge B) \cdot m(A^c \wedge B^c) > 0$) then (in both the cases) $\circ(A^c|B) = \circ(A^c|B^c) = 0$. Now, we need to look at $B|A$ and $B^c|A$, through the system (S^1) (of Theorem 3), that can be written in a compact form by referring to Pl^1 and m^1, i.e.

$$(S^1) = \begin{cases} Pl^1(A \wedge B) = Pl(B|A) \cdot Pl^1(A), \\ Pl^1(A \wedge B^c) = Pl(B^c|A) \cdot Pl^1(A), \\ m^1(A \wedge B) + m^1(A \wedge B^c) + m^1(A) = 1, \\ m^1(\cdot) \geq 0. \end{cases}$$

To second equality of condition *(aa)* of Definition 5 holds if and only if $\circ(A \wedge B) = \circ(A \wedge B^c)$, that means $Pl(B|A) \cdot Pl(B^c|A) > 0$. Then, if $Pl(B|A) \cdot Pl(B^c|A) > 0$ all the agreeing class with $Pl_{|\mathcal{D}}$ are such that $\circ(A|B) = \circ(A|B^c) = 1$ and $\circ(A^c|B) = \circ(A^c|B^c) = 0$; otherwise none agreeing class satisfies condition *(aa)*.

If $Pl(B) = 0$ (i.e. $m(A^c \wedge B^c) = 1$) then $\circ(A^c|B^c) = \circ(B^c) = 0$ and (S^1) is

$$(S^1) = \begin{cases} Pl^1(A \wedge B) = 0 \cdot Pl^1(B), \\ Pl^1(A^c \wedge B) = 1 \cdot Pl^1(B), \\ Pl^1(A \wedge B) = Pl(B|A) \cdot Pl^1(A), \\ Pl^1(A \wedge B^c) = Pl(B^c|A) \cdot Pl^1(A), \\ m^1(\cdot) \geq 0. \end{cases}$$

A solution of (S^1) is such that $m(D) = 0$ for any $D \wedge (A \wedge B) \neq \emptyset$; then when $Pl(B|A) = 0$ we need to take in consideration the following cases

- $Pl^1(A) \cdot Pl^1(B) > 0$, then it follows $Pl^1(A \wedge B^c) \cdot Pl^1(A^c \wedge B) > 0$ and $\circ(A|B) = \circ(A \wedge B) - 1 = 1, \circ(A|B^c) = 1$ and $\circ(A^c|B) = 1 - 1 = 0 = \circ(A^c|B^c)$.
- $Pl^1(A) > 0$ and $Pl^1(B) = 0$, then it follows $\circ(A|B) = \circ(A \wedge B) - 2 = 1, \circ(A|B^c) = 1$ and $\circ(A^c|B) = 2 - 2 = 0 = \circ(A^c|B^c)$.
- $Pl^1(A) = 0$ and $Pl^1(B) > 0$, then it follows $\circ(A|B) = \circ(A \wedge B) - 1 = 2, \circ(A|B^c) = 2$ and $\circ(A^c|B) = 1 - 1 = 0 = \circ(A^c|B^c)$.

On the other hand, when $Pl(B|A) > 0$, it follows from the above system $Pl^1(A) = 0$, so $Pl^1(B) = 1$ and $\circ(A^c \wedge B) = 1, \circ(A \wedge B) = 2$, while $\circ(A \wedge B^c) \geq 2$. It implies $\circ(A|B) = 1$ while $\circ(A|B^c) \geq 2$.

We can conclude this case: if $Pl(B|A) = 0$ any agreeing class of $Pl_{\mathcal{D}}$ satisfies the two equalities; while if $Pl(B|A) > 0$ no agreeing class satisfies the two equalities among the relevant zero-layers.

If $Pl(B^c) = 0$ is analogous to the previous one, just exchange B with B^c.

3. If $Pl(A^c) = 0$ is the same as 2., with A^c playing the role of A.

From the above result we get

Corollary 1 *Given a coherent conditional plausibility Pl defined on $\mathcal{F}$. If A is independent of B under Pl, then*

$$Pl(A \wedge B) = Pl(A)Pl(B).$$

It follows that the proposed notion of independence implies *cognitive independence* of Shafer [29], called also *weak independence* by Kong [27].

We have also the converse implication under suitable hypothesis, as shown in the next result.

Proposition 1 *Given a coherent conditional plausibility Pl defined on $\mathcal{F}$. If $Pl(B)$, $Pl(B^c)$, $Pl(A), Pl(A^c)$ are greater than 0, and*

$$Pl(A \wedge B) = Pl(A)Pl(B)$$
$$Pl(A \wedge B^c) = Pl(A)Pl(B^c)$$
$$Pl(A^c \wedge B) = Pl(A^c)Pl(B)$$
$$Pl(A^c \wedge B^c) = Pl(A^c)Pl(B^c)$$

then A is independent of B under $[Pl]$.

Proof: It follows directly from the definition of conditional plausibility and the properties of zero-layers.

The following example shows that the positivity condition cannot be avoided.

EXAMPLE 4 *Let A, B be two possible events and consider the assessment*
$Pl(B) = 0, Pl(B^c) = 1, Pl(A \wedge B^c) = Pl(A^c \wedge B^c) = Pl(A) = Pl(A^c) = Pl(A|B^c) = Pl(A^c|B^c) = \frac{2}{3}$,
$Pl(A|B) = Pl(A^c|B) = \frac{1}{2}$.
It is easy to show that Pl is a coherent conditional plausibility and for any atom generated by A and B, e.g. $A \wedge B$, its plausibility is equal to the product of the plausibilities of A and B, i.e.

$$Pl(A \wedge B) = Pl(A)Pl(B).$$

But, under Pl, we have that A is not independent of B.

Proposition 2 *Under any coherent conditional plausibility Pl, for any event A the statement "A is independent of itself" does not hold.*

Proof: Since by the axioms of conditional plausibilities we have that $Pl(A|A) = 1$, while $Pl(A|A^c) = Pl(\emptyset|A^c) = 0$, if follows that the statement does not hold.

The previous property (irreflexivity) is natural and essential, in fact *any event must be dependent on itself.*

Moreover, *independence implies logical independence*, as proved below. Recall that two events A and B are logically independent if all the events of the form $A^* \wedge B^*$ (where A^* stands for A or A^c) are possible, i.e. the number of relevant atoms is maximal.

Theorem 5 *Let Pl be a coherent conditional plausibility defined on $\mathcal{F}$. Given two possible events $A, B \in \mathcal{F}$, if A is independent of B under Pl, then A and B are logically independent.*

Proof: If there is a logical constraint between A and B we show that there is no agreeing class satisfying condition *(aa)*. If, for example, $A \wedge B = \emptyset$, then $Pl(A|B) = 0$ and $\circ(\emptyset) = \circ(A|B) = +\infty$; while being $A \wedge B^c = A$ a possible event $\circ(A|B^c) \leq \circ(A \wedge B^c) < +\infty$. The proof for other logical constraints follows similarly.

This is an intuitive implication: in fact *if an event is "logically" related to another, the two events must be not independent* under *any* uncertainty measure. Handling logical constraints is interesting also from a practical point of view, since in many real applications variables are suitably linked.

Remark 4 *Actually, independence under a measure assures the logical independence and this implication is guaranteed by the requirement (aa) of Definition 5. We recall that logical independence is taken into account also in [29] (as well in [2]), and it looks natural looking on Dempster rule. However, the independence notion introduced in [2] do not respect the above implication when events with degree of belief 0 are involved.*

We recall that the main difference between the approaches of [29] and [2] is that the first is referred to belief or plausibility function that are normalized (as in this paper) while in the second approach are taken in consideration also not normalized measures.

The following result characterizes independence in terms of the conditional plausibility (avoiding zero-layers).

Theorem 6 *Let A and B be two logically independent events. If a coherent conditional plausibility Pl is such that*

$$Pl(A|B) = Pl(A|B^c) = Pl(A)$$

and

$$Pl(A^c|B) = Pl(A^c|B^c) = Pl(A^c)$$

then $A \perp\!\!\!\perp B[Pl]$ if and only if one (and only one) of the following conditions holds:

1. $Pl(A) \cdot Pl(A^c) > 0$;
2. $Pl(A) = 0$ *and the coherent extension of Pl to $Pl(B)$, $Pl(B^c)$, $Pl(B|A)$, $Pl(B^c|A)$ satisfies one of the following:*
   a) $Pl(B) \cdot Pl(B^c) > 0$ *and* $Pl(B|A) \cdot Pl(B^c|A) > 0$,
   b) $Pl(B) = 0$ *and* $Pl(B|A) = 0$,
   c) $Pl(B^c) = 0$ *and* $Pl(B^c|A) = 0$;
3. $Pl(A^c) = 0$ *and the coherent extension of Pl to $Pl(B)$, $Pl(B^c)$, $Pl(B|A^c)$, $Pl(B^c|A^c)$ satisfies one of the following:*
   a) $Pl(B) \cdot Pl(B^c) > 0$ *and* $Pl(B|A^c) \cdot Pl(B^c|A^c) > 0$,
   b) $Pl(B) = 0$ *and* $Pl(B|A^c) = 0$,
   c) $Pl(B^c) = 0$ *and* $Pl(B^c|A^c) = 0$;

Proof: The items highlighted in the theorem statement follow directly by the proof of Theorem 4. In particular when $Pl(A) \cdot Pl(A^c) > 1$ is obvious because $\circ(A^*|B^*) = 0$ for all agreeing class. Moreover cases $Pl(A) = 0$ and $Pl(A^c) = 0$ correspond to the case 2. and 3. of Theorem 4 respectively: this result follows along the same proof of the previous result.

From the above result it comes out that the provided independence notion is not symmetric, and this happens when events with zero plausibility are involved. If $Pl(A), Pl(A^c), Pl(B), Pl(B^c)$ takes positive values and $A \perp\!\!\!\perp B$ under Pl, then

$$Pl(B|A) = \frac{Pl(A|B)Pl(B)}{Pl(A)} = Pl(B),$$

so going along the same computations $Pl(B|A) = Pl(B|A^c)$ and $Pl(B^c|A) = Pl(B^c|A^c) = Pl(B^c)$, which implies that also the statement $B \perp\!\!\!\perp A$ holds.

Since coherent conditional probability are a particular plausibility, and since the provided conditional independence for conditional plausibility is just a generalization of that given for conditional probability [9, 32] the fact that symmetry can fail when possible events of zero plausibility are involved is not a surprise, different examples have been given in the quoted papers to show that the lack of symmetry can be intuitive.

3.1 Independence for belief functions

By means of duality we obtain that if Pl is a coherent conditional plausibility on a set of conditional events $\mathcal{C}$ and Bel is its dual function on

$$\mathcal{C}^* = \{E|H \,:\, E^c|H \in \mathcal{C}\},$$

then Bel is a coherent conditional belief function (see Theorem 3).

Moreover, for $A|B, A|B^c, A \in \mathcal{C}$

$$Pl(A|B) = Pl(A|B^c) = Pl(A)$$

if and only if

$$Bel(A^c|B) = Bel(A^c|B^c) = Bel(A^c)$$

for $A^c|B, A^c|B^c, A^c \in \mathcal{C}$.

Then, it could seem reasonable to take A independent of B under Bel if and only if A is independent of B under the dual conditional plausibility Pl.

Recall that as shown in Section 2 a class $\mathcal{L}$ is agreeing for Bel if and only if it is agreeing also for the dual conditional plausibility Pl.

Note that this means that many properties of independence under plausibilities continue to be valid under belief functions, as for example independence implies logical independence. Moreover, also several results can be reformulated, as e.g. the characterization of independence of two possible events in terms of their belief (as done for plausibilities in Theorem 6).

Nevertheless, this notion of independence need to be studied more deeply: we need to detect better the role of zero-layers, and to exploit the relationship with the factorization property, i.e.

$$Bel(A^* \wedge B^*) = Bel(A^*)Bel(B^*).$$

Actually, the factorization has been adopt (as notion of independence) in [28] to prove under some technical hypothesis a strong law of large numbers for belief functions.

In the following example we propose a situation where, under a plausibility Pl, the event A is independent of B, but the factorization fails under the dual function of Pl.

Example 5 *Consider the following basic assignment with focal elements*

$$m_{A \wedge B} = \frac{1}{2}, m_{A \vee B} = m_\Omega = \frac{1}{4};$$

which gives rice to the following belief function $Bel(A \wedge B) = Bel(A) = Bel(B) = Bel(A \vee B^c) = Bel(A^c \vee B) = \frac{1}{2}$, $Bel(A^c \wedge B^c) = Be(A^c) = Bel(B^c) = Bel(A^c \vee B^c) = 0$, *and so to plausibility* $Pl(A \wedge B) = Pl(A) = Pl(B) = 1$ $Pl(A \wedge B^c) = Pl(A^c \wedge B) = \frac{1}{2}$ $Pl(A^c \wedge B^c) = \frac{1}{4}$.

Then, the induced conditional plausibility is such that $A \perp\!\!\!\perp B[Pl]$ *(and* $B \perp\!\!\!\perp A[Pl]$*), but* $Bel(A \wedge B) = \frac{1}{2} \neq Bel(A)Bel(B) = \frac{1}{4}$.

Thus, our notion of independence under a plausibility is stronger than *cognitive independence* [27, 29]. However, in the case of positive events it does not imply *evidential independence*, called also strong independence [29, 27], which coincides with the requirement of factorization of the belief function and its dual. However, by adding to our independence notion the factorization property with respect to the belief function, we obtain a notion stronger than evidential independence. These considerations are useful also for comparing our notion with some concepts of independence, irrelevance and non-interactivity, given in [2] (also for non necessarily normalized measures) since the notion of doxastic independence and non-interactivity coincide with evidential independence in the case of interest, i.e. for normalized measures.

3.2 Conditional independence

The notion of independence between two events given in Definition 5 can be generalized to that of conditional independence:

Definition 6 *Given a coherent conditional plausibility Pl on a set of conditional events $\mathcal{F}$ containing $\mathcal{D} = \{A^*|B^* \wedge C, A^*|C\}$, A is independent of B conditionally to C under Pl (in symbol $A \perp\!\!\!\perp B|C\ [Pl]$), if both the following conditions hold:*

i. $Pl(A|B \wedge C) = Pl(A|B^c \wedge C) = Pl(A|C)$
$Pl(A^c|B \wedge C) = Pl(A^c|B^c \wedge C) = Pl(A^c \wedge C)$,

ii. there exists an agreeing class $\mathcal{L} = \{Pl_\alpha\}$ *for the restriction of* Pl *to* $\mathcal{D}$ *such that*

$$\circ(A|B \wedge C) = \circ(A|B^c \wedge C) \text{ and}$$
$$\circ(A^c|B \wedge C) = \circ(A^c|B^c \wedge C).$$

Considerations similar to the unconditional case can be done: when $Pl(A|B \wedge C)$ and $Pl(A^c|B \wedge C)$ are both positive, then both equalities in condition *ii.* are trivially (as 0=0) satisfied. While in the other two cases (i.e. $Pl(A|B \wedge C) = 0$ or $Pl(A^c|B \wedge C) = 0$) the equality *i.* is not enough to assure independence, so it is "reinforced" by the requirement that also their relevant zero-layers must be equal.

Remark 5 *If the events* A *and* C *(or* A^c *and* C*) are incompatible, then* A *is independent of any event* B *given* C *whenever* $\emptyset \neq B \wedge C \neq C$*. This conclusion is natural since the plausibility* 0 *(or* 1*) of* $A|C$ *cannot be changed by assuming the occurrence of* B*.*

Actually, even if the restriction of Pl to $\mathcal{D}$ admits more than one agreeing class, we can prove along the line of Theorem 4, that condition *ii.* of Definition 6 holds either for all agreeing classes or for none of them.

Going on the same line of the proofs given for Theorem 4 and Theorem 6, we can characterize conditional independence in terms of plausibilities: it would be a simply generalization of Theorem 6.

4 Summary and Conclusions

In this paper, we look to conditional plausibility and belief from a more general point of view. In particular we are able to handle events with null measure. In this framework we provide a definition of independence and we give a characterization of it, we study its main properties, which allow us to compare our definition with other given in literature (in particular with respect to the notions introduced in [2, 3, 27, 29]). We recall that our notion of independence for plausibility is in the same line of that studied in [9, 11, 24, 32] for probability and possibility.

Through different examples we explain also the reason for taking exactly the provided definition, our choice has been guided mainly by two main reasons: to get a natural condition overcoming critical aspects and to get a suitable factorization of the joint plausibility distribution.

We show that the provided independence notion is not necessarily symmetric, then to represent such statements we need to refer to some not necessarily symmetric separation criterion such as that proposed in [33]. An open problem consists into looking for the representability of the set of independence statements induced by a conditional plausibility (belief) by means of a directed or undirected graph by testing which properties among the graphoid ones are satisfied: this would allow to compare our definition also with other independence notions given in the context of other uncertainty formalisms.

References

[1] N. Ben Amor, K. Mellouli, S. Benferhat, D. Dubois, H. Prade. A theoretical framework for possibilistic independence in a weakly ordered setting. Int. Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 10(2), 117–155, 2002.

[2] B. Ben Yaghlane, P. Smets, K. Mellouli. Belief function independence: I. The marginal case *Int. Journal of Approximate Reasoning*, 29, 47-70, 2002.

[3] B. Ben Yaghlane, P. Smets, K. Mellouli. Belief function independence: I. The conditional case *Int. Journal of Approximate Reasoning*, 31, 31-75, 2002.

[4] B. Bouchon-Meunier, G. Coletti, C. Marsala. Independence and Possibilistic Conditioning. *Annals of Mathematics and Artificial Intelligence*, 35, . 107-124, 2002.

[5] L.M. Campos, J.F. Huete. Independence concepts in upper and lower probabilities. In: *Uncertainty in Intelligent Systems* (Eds. Bouchon-Meunier et al.), 85-96, 1993.

[6] G. Coletti, M. Mastroleo Conditional belief functions: a comparison among different definitions. *Proceeding of 7th Workshop on Uncertainty Processing (WUPES 2007)*, 2006.

[7] G. Coletti, R. Scozzafava. From conditional events to conditional measures: a new axiomatic approach. *Annals of Mathematics and Artificial Intelligence* 32, 373–392, 2001.

[8] G. Coletti, R. Scozzafava. Stochastic Independence in a Coherent Setting. *Annals of Mathematics and Artificial Intelligence* 35, 151176, 2002.

[9] G. Coletti, R. Scozzafava. Probabilistic Logic in a Coherent Setting (Trends in Logic, n.15). Dordrecht: Kluwer: 2002.

[10] G. Coletti, R. Scozzafava. Toward a General Theory of Conditional Beliefs. *International Journal of Intelligent Systems* 21, 229-259, 2006.

[11] G. Coletti, B. Vantaggi. Possibility theory: conditional independence. *Fuzzy Sets and Systems*, 157(11), 1491–1513, 2006.

[12] G. Coletti, B. Vantaggi. A view on conditional measures through local representability of binary relations. Submitted in *International Journal of Approximate Reasoning*.

[13] F.G. Cozman. Irrelevance and independence axioms in quasi-Bayesian theory, in: T. Hunter, S. Parsons (Eds.), *Proc. of ECSQARU99, Lecture Notes in AI 1638*, Springer-Verlag, Berlin, 1999, pp. 128136, 2006.

[14] R.T. Cox. Probability, frequency and reasonable expectation. *American Journal of Physics*, 14 (1), 1–13, 1946.

[15] A.P. Dawid Conditional independence in statistical theory. *Journal of Royal Statistical Society B*, 41, 15-31, 1979.

[16] B. de Finetti. Sull'impostazione assiomatica del calcolo delle probabilità. *Annali Univ. Trieste*, 19, 3–55 1949 - Engl. transl.: Ch.5 in *Probability, Induction, Statistics*, Wiley, London, 1972.

[17] B. de Finetti. Les probabilitès nulles. *Bull. Sci. Math.*, 60, 175–288, 1936.

[18] A.P. Dempster. A generalizatin of Bayesian Inference. *The Royal Stat. Soc. B*, 50, 205–247, 1968.

[19] T. Denoeux, P. Smets, Classification using Belief Functions: the Relationship between the Case-based and Model-based Approaches, To appear in *IEEE Transactions on Systems, Man and Cybernetics B*, 2006.

[20] F.T. de Dombal, F. Gremy. *Decision Making and Medical Care*. North-Holland, 1976.

[21] J.R. Hill. Comment on Graphical models. *Statistical Science* 8, 258-261, 1993.

[22] D. Dubois, L. Fariñas del Cerro, A. Herzig, H. Prade. An ordinal view of independence with application to plausible reasoning. *Proc. of the 10th Conf. on Uncertainty in Artificial Intelligence* (R. Lopez de Mantaras, D. Poole, eds.), Seattle, WA, 195-203 1994.

[23] Dubois, H. Prade, Possibility Theory New York, Plenum Press, 1988.

[24] L. Ferracuti, B. Vantaggi. Independence and conditional possibilities for strictly monotone triangular norms. *International Journal of Intelligent Systems*, 21(3), 299-323, 2006.

[25] J.Y. Halpern. A counterexample to theorems of Cox and Fine. *J. of Artificial Intelligence Research*, 10, 67–85, 1999.

[26] J.Y. Jaffray. Bayesian updatingand Belief functions. *IEEE Trans. on Systems, man and Cybernetic*, 5, 1144–1152, 1992.

[27] C.T.A. Kong. A belief function generalization of Gibbs Ensemble. Tec. Report S-122 Harvard University, 1988.

[28] F. Maccheroni, M. Marinacci. A strong law of large numbers for capacities. *Annals of Probabilities*, 33(3), 1171–1178, 2005.

[29] G. Shafer. *A Mathematical Theory of Evidence*, Princeton University Press, Princeton, NJ, 1976.

[30] W. Spohn. On the Properties of Conditional Independence, in: Humphreys, Suppes, (Eds), Scientific Philosopher 1: Probability and Probabilistic Causality, Kluwer, Dordrecht, 173–194.

[31] M. Studeny. Formal properties of conditional independence in different calculi of AI, in: Clarke, Kruse, Moral (Eds), Symbolic and Quantitative Approaches to Reasoning and Uncertainty, Berlin, Springer-Verlag, 341–348, 1993.

[32] B. Vantaggi. Conditional independence in a finite coherent setting. *Annals of Mathematics and Artificial Intelligence*, 32, 287-314, 2001.

[33] Vantaggi B. The L-separation criterion for description of cs-independence models. *International Journal of Approximate Reasoning*, 29, 291-316, 2002.

[34] B. Vantaggi. Conditional independence structures and graphical models. *Int. Journal of Uncertainty, Fuzziness and Knowledge-based Systems* 11(5), 545-571, 2003.

[35] J. Vejnarova. Conditional Independence Relations in Possibility Theory. *International Journal of Uncertainty, Fuzziness and Knowledge-based Systems* 8(3), 253-269, 2000.

[36] P. Walley. Belief function representations of statistical evidence. *Annals of Statistics*, 4, 1439–1465, 1987.