Independence for Sets of Full Conditional Probabilities, Sets of Lexicographic Probabilities, and Sets of Desirable Gambles

Fabio Gagliardi Cozman
Escola Politecnica - Universidade de Sao Paulo
Av. Prof. Mello Moraes 2231 - Cidade Universitaria, Sao Paulo, SP - Brazil

Abstract

In this paper we examine concepts of independence for sets of full conditional probabilities; that is, for sets of set-functions where conditional probability is the primitive concept, and where conditioning can be considered on events of probability zero. We also discuss the related issue of independence for (sets of) lexicographic probabilities and for sets of desirable gambles.

Keywords. Sets of probability measures, full conditional probabilities, lexicographic probability, sets of desirable gambles, independence concepts, graphoids.

1 Introduction

This paper examines concepts of independence for sets of full conditional probabilities and related models. We study the behavior of several concepts of independence in the literature, and propose a number of possible additional concepts. The results should be of interest to anyone concerned with representations of uncertainty that allow indeterminacy and imprecision in probability values, and that allow conditioning on every nonempty event.

The motivation for this paper is the following.

The use of a single standard probability measure fails to encode indeterminacy and imprecision about probability values. Belief functions, interval-valued probability, and sets of probability measures have been proposed to handle such indeterminacy and imprecision. It is not obvious how to generalize the concept of stochastic independence when one deals with sets of probability measures; accordingly, there have been many proposed concepts of independence in the literature.

Another problem with standard probability measures is that they do not handle conditioning on events of probability zero; that is, if $P(B) = 0$, then $P(A|B)$ does not exist, regardless of the event A. Indeed, standard conditional probability is merely a derived, incompletely specified concept, while one might argue that conditional probability should be the primitive object of interest. Full conditional probabilities offer an account of conditional probability as primitive objects that can be specified even if conditioning events have probability zero. As standard stochastic independence is quite weak when applied to full conditional probabilities, there have been several proposals for concepts of independence that are appropriate for a single full conditional probability.

However, there is still much to be understood about concepts of independence for *sets* of full conditional probabilities. This paper tries to partially fill this gap, by examining a number of concepts of independence and deriving their graphoid properties (these properties are often taken as abstract properties that any "sensible" concept of independence should satisfy). We also discuss concepts of independence for (sets of) lexicographic probabilities and sets of desirable gambles, as they share several features with full conditional probabilities.

Section 2 describes existing and novel concepts of independence for credal sets and full conditional probabilities. It does not seem that a similar analysis can be found in the literature. Section 3 examines a number of new concepts of independence for sets of full conditional probabilities. Section 4 then examines concepts of independence that resort to lexicographic probabilities and to sets of desirable gambles.

2 Concepts of independence

We assume throughout that the possibility space Ω is finite, so there are no issues of measurability. Throughout the paper we use W, X, Y and Z to denote random variables. Then w denotes a possible value of W, x denotes a possible value of X, y denotes a possible value of Y, z denotes a possible value of Z. And $\{x\}$ denotes the event $\{\omega \in \Omega : X(\omega) = x\}$; likewise for $\{w\}$, $\{y\}$ and $\{z\}$. The letters A and C will always denote nonempty events in the algebra generated by X. Likewise, the letters B and D will always denote nonempty events in the algebra generated by Y. The letter f will always denote a function of X, and the letter g will always denote a function of Y.

The intersection of events G and H is written either as GH or as G, H. When the event $\{x\}$ appears in an intersection, we remove braces whenever possible; for instance, xG denotes the event $\{x\} \cap G$. Sometimes we add braces to enhance clarity; for instance, we may write $\{y, z\}$ instead of simply y, z.

Finally, when w, x, y, z appear in expressions, they are universally quantified unless explicitly noted. Likewise, when functions f and g appear in expressions, they are universally quantified unless explicitly noted.

Conditional stochastic independence of random variables X and Y given random variable Z obtains when $P(x, y|z) = P(x|z)\, P(y|z)$ whenever $P(z) > 0$.

Throughout, if Z is any constant function, we remove the expression "given Z" and in that case we have "unconditional" independence of X and Y (for any concept of independence of interest). Often we just write "independence" to mean both conditional and unconditional independence.

Concepts of independence can be evaluated by their graphoid properties [14, 34]. For any three-place relation $(\cdot \perp\!\!\!\perp \cdot \,|\, \cdot)$, we are interested in the following properties, all of them satisfied by stochastic independence:

Symmetry: $(X \perp\!\!\!\perp Y \,|\, Z) \Rightarrow (Y \perp\!\!\!\perp X \,|\, Z)$

Redundancy: $(X \perp\!\!\!\perp Y \,|\, X)$

Decomposition: $(X \perp\!\!\!\perp (W, Y) \,|\, Z) \Rightarrow (X \perp\!\!\!\perp Y \,|\, Z)$

Weak union: $(X \perp\!\!\!\perp (W, Y) \,|\, Z) \Rightarrow (X \perp\!\!\!\perp Y \,|\, (W, Z))$

Contraction:
$(X \perp\!\!\!\perp Y \,|\, Z) \wedge (X \perp\!\!\!\perp W \,|\, (Y, Z)) \Rightarrow (X \perp\!\!\!\perp (W, Y) \,|\, Z)$.

2.1 Independence for sets of standard probability measures

A set of standard (Kolmogorovian-style) probability measures, not assumed to be closed and convex, is referred to as a *credal set*. Denote by $K(X)$ the set of probability distributions for variable X. Given a function $f(X)$, its lower and upper expectations are, respectively $\underline{E}[f(X)] = \inf_{P \in K} E_P[f(X)]$ and $\overline{E}[f(X)] = \sup_{P \in K} E_P[f(X)]$, where $E_P[f(X)]$ is the expectation of $f(X)$ with respect to P. Similarly, given an event A, its lower and upper probabilities are, respectively $\underline{P}(A) = \inf_{P \in K} P(A)$ and $\overline{P}(A) = \sup_{P \in K} P(A)$.

Given a credal set $K(X)$, we define the conditional credal set

$$K(X|A) = \{P(\cdot|A) : P \in K(X)\} \quad \text{if } \underline{P}(A) > 0;$$

otherwise, $K(X|A)$ is left undefined [21]. Another option is to define a conditional credal set that focuses on those probability measures that assign positive probability to A:

$$K^{>}(X|A) = \{P(\cdot|A) : P \in K(X) \text{ and } P(A) > 0\} \quad \text{if } \overline{P}(A) > 0; \tag{1}$$

otherwise $K^{>}(X|A)$ is left undefined [44, 45]. Obviously, if $\underline{P}(A) > 0$, then $K(X|A) = K^{>}(X|A)$. The set $K^{>}(X|A)$ is convex when $K(X)$ is convex, but it may be open even when $K(X)$ is closed. We define $\underline{E}^{>}[f(X)|A] = \inf_{P(\cdot|A) \in K^{>}(X|A)} E_P[f(X)|A]$ and $\overline{E}^{>}[f(X)|A] = \sup_{P(\cdot|A) \in K^{>}(X|A)} E_P[f(X)|A]$.

For a moment, assume that all lower probabilities are positive.

Following Levi [29], say that Y is *confirmationally irrelevant* to X given Z when

$$K(X|y, z) = K(X|z)\,. \tag{2}$$

Walley has proposed a similar concept [41, 42]: Y is *epistemically irrelevant* to X given Z when

$$\underline{E}[f(X)|y, z] = \underline{E}[f(X)|z] \tag{3}$$

(recall our conventions: by implicit quantification, this equality is required for all f, for all y, z).

Both confirmational and epistemic irrelevance fail Symmetry. Walley's clever solution, borrowed from the work of Keynes, was to "symmetrize" irrelevance to obtain *epistemic independence*: X and Y are epistemically independent given Z when X is epistemically irrelevant to Y given Z and Y is epistemically irrelevant to X given Z [42]. Take *confirmational independence* to be a likewise symmetrized version of confirmational irrelevance.

If all credal sets are closed and convex, then confirmational and epistemic independence are equivalent. Now even if all lower probabilities are positive and all credal sets are closed and convex, epistemic independence (and confirmational independence) fails Contraction [7]. And if credal sets are not required to be convex, then confirmational independence fails Decomposition, Weak Union and Contraction even when all lower probabilities are positive [9].

Matters become more complicated if lower probabilites are allowed to be zero. Suppose first that Y is taken to be confirmationally irrelevant to X if

$$K(X|y,z) = K(X|z) \text{ whenever } \underline{P}(y,z) > 0.$$

We are surely flirting with disaster here, because it is not difficult to have a variable Z such that every value of Z has zero lower probability, and yet $K(Z)$ is not a vacuous credal set (that is, it does not contain every possible distribution for Z). Now given such a variable Z, every two other variables are confirmationally independent! This is not reasonable.

The other path to handle events of zero lower probability within confirmational independence is to say that Y is confirmationally irrelevant to X given Z when

$$K^{>}(X|y,z) = K(X|z) \text{ whenever } \overline{P}(y,z) > 0. \quad (4)$$

The symmetrized concept of independence fails Decomposition, Weak Union and Contraction (as noted before, these properties fail even when all lower probabilities are positive [9]).

Another possibility is to define epistemic irrelevance of Y to X given Z by requiring:

$$\underline{E}^{>}[f(X)|y,z] = \underline{E}[f(X)|z] \text{ whenever } \overline{P}(y,z) > 0. \quad (5)$$

The resulting symmetrized concept of independence fails Contraction (as noted before, this property fails even when all lower probabilities are positive [7]). It is an open question whether Decomposition and Weak Union hold when Expression (5) is used to define independence; Decomposition and Weak Union hold for epistemic independence when all lower probabilities are positive [12].

Note: Expressions (4) and (5) impose different constraints, as $K^{>}(X|A)$ may be open even when $K(X)$ is closed.

Yet another path has been followed by de Campos and Moral [15]: they say Y is *type-5* irrelevant to X if

$$K^{>}(X|B) = K(X) \text{ whenever } \overline{P}(B) > 0$$

(recall: B is an event in the algebra generated by Y). Accordingly, say that Y is *type-5 irrelevant* to X given Z if

$$K^{>}(X|B,z) = K(X|z) \text{ whenever } \overline{P}(B,z) > 0.$$

Now we might also modify epistemic irrelevance, and say that Y is *type-5 epistemically irrelevant* to X given Z if

$$\underline{E}^{>}[f(X)|B,z] = \underline{E}[f(X)|z] \text{ whenever } \overline{P}(B,z) > 0.$$

And we can symmetrize type-5 irrelevance and type-5 epistemic irrelevance to obtain corresponding concepts of independence. Now, Contraction fails for type-5 independence and for type-5 epistemic independence (Contraction fails already when all lower probabilities are positive [7]). It is an open question whether Weak Union holds for these concepts of independence. As for Decomposition:

Proposition 1 *Both type-5 independence and type-5 epistemic independence satisfy Decomposition.*

Proof. Assume X and (W,Y) are type-5 independent given Z. Then $K(Y|A,z) = K(Y|z)$ by marginalization, and $K(X|B,z) = K(X|z)$ because any B belongs to the algebra generated by (W,Y). Likewise, assume type-5 epistemic independence holds for X and (W,Y). Then $\underline{E}[g(Y)|A,z] = \underline{E}[g(Y)|z]$ because any function of Y is a function of (W,Y), and $\underline{E}[f(X)|B,z] = \underline{E}[f(X)|z]$. $\square$

Type-5 irrelevance may seem very attractive at first, but the following example, due to de Campos and Moral [15], displays rather weird behavior when lower probabilities are zero. Take binary variables X and Y, and $K(X,Y)$ with two distributions, one that assigns probability one to (x_0,y_0) and another that assigns probability one to (x_1,y_1) (if $K(X,Y)$ must be convex, take the convex hull of these two distributions). Both distributions satisfy stochastic independence, but X and Y fail to be type-5 independent! In general, type-5 independence may fail even when all elements of the credal set $K(X,Y)$ factorize.

This discussion suggests that concepts of independence for credal sets must handle conditioning carefully. We now describe a few concepts of independence that require no discussion about conditioning.

Strong independence was also proposed by Levi [29], initially with the name *strong confirmational irrelevance*: X and Y are strongly independent when $K(X,Y)$ is the convex hull of a set of probability measures that satisfy stochastic independence. Strong independence is an attempt to stay close to stochastic independence while assuming convexity (given that imposing stochastic independence over a set of probability measures may generate a nonconvex set of measures). Strong independence can be derived from assumptions of infinite exchangeability [9] or finite exchangeability together with epistemic independence [16]. Note that strong independence, and slight variants of it, have received several names in the literature, such as *type-1 product*, *type-2 product*, *type-2 independence*, *independence in the selection*, *repetition independence* [9].

Complete independence abandons convexity and imposes stochastic independence directly: X and Y are completely independent when every joint distribution in $K(X,Y)$ satisfies stochastic independence [9]. Complete independence satisfies all graphoid properties previously mentioned.

The last notable concept of independence we mention for credal sets is due to Kuznetsov [28]: X and Y are Kuznetsov-independent if

$$\mathbb{E}[f(X)g(Y)] = \mathbb{E}[f(X)] \boxtimes \mathbb{E}[g(Y)]$$

for all functions $f(X)$ and $g(Y)$, where $\mathbb{E}[\cdot]$ denotes the interval from lower to upper expectations, and $\boxtimes$ denotes interval multiplication. Kuznetsov-independence satisfies Symmetry, Redundancy and Decomposition; it fails Contraction even when all probabilities are positive [8], and it is an open question whether it satisfies Weak Union or not.

2.2 Independence for full conditional probabilities

A full conditional probability [20] $P : \mathcal{B} \times (\mathcal{B}\backslash\emptyset) \to \Re$, where $\mathcal{B}$ is a Boolean algebra, is a two-place set-function such that for every event $H \neq \emptyset$:
(1) $P(H|H) = 1$;
(2) $P(G|H) \geq 0$ for all G;
(3) $P(G_1 \cup G_2|H) = P(G_1|H) + P(G_2|H)$ whenever $G_1 \cap G_2 = \emptyset$;
(4) $P(G_1, G_2|H) = P(G_1|G_2, H) \times P(G_2|H)$ whenever $G_2 H \neq \emptyset$.

This fourth axiom is often stated as $P(G_1|H) = P(G_1|G_2)\, P(G_2|H)$ when $G_1 \subseteq G_2 \subseteq H$ and $G_2 \neq \emptyset$ [13, Section 2].

Define the "unconditional" probability $P(G)$ of an event G to be $P(G|\Omega)$. That is, whenever the conditioning event H is equal to Ω, we suppress it and write the "unconditional" probability $P(G)$.

There are other names for full conditional probabilities in the literature, such as *conditional probabilities* [27] and *complete conditional probability systems* [33]. We simplify to *full probability* whenever possible. Full probabilities have found applications in several fields, notably economy, philosophy, and statistics [5, 19, 26, 30, 32, 35, 38].

We can partition Ω into events $L_0, \dots, L_K$ as follows. First, take L_0 to be the set of elements of Ω that have positive unconditional probability. Then take L_1 to be the set of elements of Ω that have positive probability conditional on $\Omega\backslash L_0$. And then take L_i, for $i \in \{2, \dots, K\}$, to be the set of elements of Ω that have positive probability conditional on $\Omega\backslash \cup_{j=0}^{i-1} L_j$. The events L_i are called the *layers* of the full probability. Note that some authors use a different terminology, using instead the sequence $\cup_{j=i}^{K} L_j$ rather than L_i [5, 27].

	y_0	y_1
x_0	$\lfloor 1 \rfloor_0$	$\lfloor 1-\alpha \rfloor_1$
x_1	$\lfloor \alpha \rfloor_1$	$\lfloor 1 \rfloor_2$

	y_0	y_1
x_0	$\lfloor 1 \rfloor_0$	$\lfloor 1 \rfloor_i$
x_1	$\lfloor 1 \rfloor_j$	$\lfloor 1 \rfloor_3$

Table 1: Joint full distributions for binary variables X and Y. The right table stands for two full distributions: one for $i = 1, j = 2$; another for $i = 2, j = 1$.

Any full probability can be represented by a sequence of probability measures $P_0, \dots, P_K$, where P_i is positive over L_i. This useful result that has been derived by several authors [3, 5, 23, 27].

For nonempty G, denote by L_G the first layer such that $P(G|L_G) > 0$, and refer to it as the *layer of* G. We then have $P(G|H) = P(G|H \cap L_H)$ [2, Lemma 2.1a].

We often write $\lfloor \alpha \rfloor_i$ to denote a probability value α that belongs to the ith layer L_i. Table 1 shows three full distributions using this compact notation.

Given a full probability and a nonempty event H, the two-place function $P(\cdot|\cdot \cap H)$ is also a full probability from which a partition of H consisting of layers $L_{0|H}, L_{1|H}, \dots, L_{K|H}$ can be built. Given an event G such that $G \cap H \neq \emptyset$, denote by $L_{G|H}$ the first layer of $P(\cdot|\cdot \cap H)$ such that $P\big(G|L_{G|H}\big) > 0$.

For a nonempty event G, the index i of the first layer L_i of the full probability P such that $P(G|L_i) > 0$ is the *layer number* of G. Layer numbers have been studied by Coletti and Scozzafava [5], who refer to them as *zero-layers*. The layer number of G is denoted by $\circ(G)$. Inspired by Coletti and Scozzafava [5], we define the layer number of G given nonempty H as $\circ(G|H) = \circ(G \cap H) - \circ(H)$, and we adopt $\circ(\emptyset) = \infty$.

Now consider concepts of independence for full probabilities.

Stochastic independence satisfies all graphoid properties we have mentioned previously, when applied to full probabilities. Unfortunately, it may happen that X and Y are stochastically independent and yet $P(A|B) \neq P(A)$ when $P(B) = 0$. Table 2 shows an extreme example. To avoid this embarrassment, more stringent notions of independence have been proposed for full probabilities [3, 5, 23, 39].

Say that Y is *epistemically irrelevant* to X given Z if $P(A|y,z) = P(A|z)$ whenever $\{y,z\} \neq \emptyset$, and then say that X and Y are *epistemically independent* given Z if X is epistemically irrelevant to Y given Z and vice-versa. Epistemic independence satisfies Sym-

	y_0	y_1
x_0	$\lfloor 1 \rfloor_0$	$\lfloor 1 \rfloor_3$
x_1	$\lfloor 1 \rfloor_1$	$\lfloor 1 \rfloor_2$

Table 2: Joint full distributions for stochastically independent binary variables, where $P(x_0) = 1 \neq 0 = P(x_0|y_1)$.

	$w_0 y_0$	$w_1 y_0$	$w_0 y_1$	$w_1 y_1$
x_0	$\lfloor \alpha \rfloor_0$	$\lfloor \beta \rfloor_2$	$\lfloor 1-\alpha \rfloor_0$	$\lfloor 1-\beta \rfloor_2$
x_1	$\lfloor \alpha \rfloor_1$	$\lfloor \gamma \rfloor_3$	$\lfloor 1-\alpha \rfloor_1$	$\lfloor 1-\gamma \rfloor_3$

Table 3: Full distribution for W, X, Y, with distinct $\alpha \in (0,1)$, $\beta \in (0,1)$, $\gamma \in (0,1)$.

metry, Redundancy, Decomposition and Contraction, but it fails Weak Union [11, Proposition 4.2]. The full distribution in Table 3 displays failure of Weak Union for epistemic independence.

As proposed by Hammond [23], say that Y is *h-irrelevant* to X given Z when

$$P(A|B,C,z) = P(A|C,z) \text{ whenever } \{B,C,z\} \neq \emptyset,$$

and say that X and Y are *h-independent* given Z when X is h-irrelevant to Y given Z and vice-versa (recall our conventions: this equality must hold for every A and C in the algebra generated by X, and for every B in the algebra generated by Y).

If X and Y are h-independent given Z, then

$$\begin{aligned} P(A,B|C,D,z) &= P(A|C,z)\,P(B|D,z) \\ &\quad \text{whenever } \{C,D,z\} \neq \emptyset. \end{aligned}$$

H-independence satisfies Symmetry, Redundancy, Decomposition and Weak Union, but it fails Contraction [11, Theorem 5.4]. The full distribution in Table 3 displays failure of Contraction for h-independence.

Coletti and Scozzafava [5] have proposed conditions on zero-layers to characterize independence. Say that event H is *cs-irrelevant* to event G, where $H \neq \emptyset \neq H^c$, if $P(G|H) = P(G|H^c)$, $\circ(G|H) = \circ(G|H^c)$, and $\circ(G^c|H) = \circ(G^c|H^c)$. To understand the motivation for these conditions on layer numbers, suppose that GH, GH^c, G^cH are nonempty, but $G^cH^c = \emptyset$. Hence observation of H^c does provide information about G. However, the indicator functions of G and H can be epistemically/h-independent! Coletti and Scozzafava eliminate such difficulties using their conditions on layer numbers; other authors, such as Hammond [23] and Battigalli [2], explicitly require the possibility space to be the product of the possibility spaces for each of the variables.

Vantaggi [39, 40] has extended Coletti and Scozzafava conditions to independence of variables. Say that Y is *cs-irrelevant* to X given Z when event $\{y\}$ is cs-irrelevant to event $\{x\}$ given event $\{z\}$, whenever $\{y,z\} \neq \emptyset \neq \{\{y\}^c, z\}$ [39, Definition 7.3]. Call the symmetrized concept *cs-independence* of X and Y given Z. Besides Symmetry, cs-independence satisfies Redundancy, Decomposition and Contraction, and it fails Weak Union [39, Section 9].

The conditions on layer numbers imposed by cs-independence can be written as [11, Corollary 4.11]:

$$\circ(x,y|z) = \circ(x|z) + \circ(y|z) \quad \text{for } \{z\} \neq \emptyset. \tag{6}$$

Condition (6) can be used to generate additional concepts of independence. For instance, say that Y is *fully irrelevant* to X given Z if Y is h-irrelevant to X given Z and if they satisfy Condition (6); say that X and Y are *fully independent* given Z if they are h-independent given Z and satisfy Condition (6) [11].

Full independence satisfies Symmetry, Redundancy, Decomposition and Weak Union, but it fails Contraction [11, Theorem 5.7]. Table 3 displays failure of Contraction for full independence.

A different concept of independence has been proposed by Kohlberg and Reny [26], essentially as follows. Say that X and Y are kr-independent given Z when both:

- if $\{x,z\} \neq \emptyset$ and $\{y,z\} \neq \emptyset$, then $\{x,y,z\} \neq \emptyset$;
- if, whenever conditioning events are nonempty,
  $$\frac{P\big(x,y|L_{x,y|z} \cup L_{x',y'|z}\big)}{P\big(x',y'|L_{x,y|z} \cup L_{x',y'|z}\big)} = \lim_{n\to\infty} \frac{P_n(x|z)P_n(y|z)}{P_n(x'|z)P_n(y'|z)}$$
  for some sequence of product probability measures $P_n(\cdot|z)$.

Relatively little is known about kr-independence; we only note that it satisfies Symmetry, Redundancy, Decomposition and Weak Union, and it fails Contraction as can be seen in Table 3 [10, Theorem 1].

We now introduce a new concept of independence for full probabilities where we require factorization across layers of the full probability [10]. Consider:

Definition 1 *X and Y are* layer independent *given Z if, for each layer L_i of the underlying full probability P, and each z such that $\{L_i, z\} \neq \emptyset$, we have both*

$$P(x,y|L_i,z) = P(x|L_i,z)\,P(y|L_i,z)\,,$$

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

This concept of independence satisfies Symmetry, Redundancy, Decomposition, Weak Union and Contraction; in fact, this seems to be the only known concept of independence for full probabilities that satisfies all these five properties.

We conclude this section by commenting on an aspect of full probabilities that has not received the deserved attention so far; namely, failure of uniqueness (some comments about it appear in the work of Battigalli [1] and Kohlberg and Reny [26]). The issue is this. Suppose one is given marginal probabilities $P(x_0) = P(y_0) = 1$ for binary variables X and Y. Now *every* full distribution in Table 1 (for *every* $\alpha \in (0, 1)$) satisfies these marginal assessments and epistemic/h-/cs-/full/kr-independence; moreover, the two full distributions encoded by the right table satisfy layer independence. In general, one cannot uniquely determine a single full probability by specifying marginal assessments and judgments of independence. Once assessments are to be combined with existing concepts of independence, one must be prepared to consider a set of joint full probabilities that satisfies all constraints.

3 Full credal sets and independence

We now focus on sets of full probabilities, and investigate the graphoid properties of several concepts of independence. We refer to such sets as *full credal sets*; we do not assume the sets to be convex and closed.

As already noted, a concept of independence that relies on product factorizations is too weak in the context of full probabilities. Indeed we have that Kuznetsov, strong, complete and type-5 independence declare X and Y independent for the full credal set containing only the full distribution in Table 2.

Complete independence can be adapted to full credal sets as follows. Define *elementwise epistemic/h-/cs-/full/kr-/layer independence* of X and Y given Z to hold when every element of the full credal set $K(X, Y|z)$ satisfies respectively epistemic/h-/cs-/full/kr-/layer independence whenever $\{z\} \neq \emptyset$. We note that Coletti and Scozzafava's concept of independence for lower probabilities [4, Definition 6], extended to variables by Vantaggi [40, Definition 7], is quite similar to elementwise cs-independence.

Given the results mentioned in the previous section:

Proposition 2 *Elementwise epistemic/cs-independence satisfy Symmetry, Redundancy, Decomposition and Contraction (and fail Weak Union). Elementwise h-/full/kr-independence satisfy Symmetry, Redundancy, Decomposition and Weak Union (and fail Contraction). Elementwise layer independence satisfies Symmetry, Redundancy, Decomposition, Weak Union and Contraction.*

A challenge that merits future work is to justify these concepts of independence from behavioral or decision-theoretic arguments. Even though complete independence has an intuitive justification using choice functions [9, 37], the interaction between choice functions and full probabilities is yet to be explored.

Consider now confirmational and epistemic independence as defined in Section 2.1, but applied to full credal sets. The resulting concepts were originally proposed by Levi [29] and by Walley [42] within theories that adopt full probabilities.

Confirmational independence fails Decomposition, Weak Union and Contraction when applied to general full credal sets (even when all lower probabilities are positive [9]).

Epistemic independence fails Decomposition and Weak Union when applied to full credal sets [12], as can be seen in Example 1, and fails Contraction even when all lower probabilities are positive [7].

Example 1 Consider a full credal set with the two distributions depicted in Table 4, where $\alpha \in (0, 1/2)$. We have $P(w_0) \in [\alpha, 1-\alpha]$ and $P(w_0|x, y) \in [\alpha, 1-\alpha]$ for all possible x, y: (X, Y) is epistemically irrelevant to W. The reader can verify that both distributions yield identical values of $P(x, y|w)$ and $P(x, y)$ such that $P(x, y|w) = P(x, y)$, for all possible (x, y, z). Hence W is epistemically irrelevant to (X, Y). Thus we have epistemic independence of W and (X, Y). However, $P(w_0|x_1) = 1/2$; consequently, X is not epistemically irrelevant to W (Decomposition fails), and Y is not epistemically irrelevant to W given X (Weak Union fails). □

So, at least from the point of view of graphoid properties, both confirmational and epistemic independence fare rather poorly.

Note that the motivation behind confirmational/epistemic irrelevance of Y to X is that observation of Y does not change beliefs about X. However, for a full probability the beliefs about X are encoded not just by expectations $E[f(X)]$ but rather by conditional expectations $E[f(X)|A]$ for events A in the algebra generated by X. This is indeed the rationale behind h-independence; for this reason, the combination of h-independence and full credal sets seems very attractive.

Consider then adapting h-independence to full credal sets as follows:

P_1	w_0y_0	w_0y_1	w_1y_0	w_1y_1
x_0	$\lfloor \frac{\alpha}{2} \rfloor_0$	$\lfloor \frac{\alpha}{2} \rfloor_0$	$\lfloor \frac{1-\alpha}{2} \rfloor_0$	$\lfloor \frac{1-\alpha}{2} \rfloor_0$
x_1	$\lfloor \frac{\alpha}{2} \rfloor_1$	$\lfloor \frac{1-\alpha}{2} \rfloor_1$	$\lfloor \frac{1-\alpha}{2} \rfloor_1$	$\lfloor \frac{\alpha}{2} \rfloor_1$

P_2	w_0y_0	w_0y_1	w_1y_0	w_1y_1
x_0	$\lfloor \frac{1-\alpha}{2} \rfloor_0$	$\lfloor \frac{1-\alpha}{2} \rfloor_0$	$\lfloor \frac{\alpha}{2} \rfloor_0$	$\lfloor \frac{\alpha}{2} \rfloor_0$
x_1	$\lfloor \frac{1-\alpha}{2} \rfloor_1$	$\lfloor \frac{\alpha}{2} \rfloor_1$	$\lfloor \frac{\alpha}{2} \rfloor_1$	$\lfloor \frac{1-\alpha}{2} \rfloor_1$

Table 4: Extreme points of the full credal set in Example 1.

	y_0	y_1
x_0	$\lfloor \alpha \rfloor_0$	$\lfloor 1-\alpha \rfloor_0$
x_1	$\lfloor \alpha \rfloor_1$	$\lfloor 1-\alpha \rfloor_1$

Table 5: Marginal probabilities from Table 3.

	y_0	y_1
x_0	$\lfloor \alpha \rfloor_0$, $\lfloor \beta \rfloor_2$	$\lfloor 1-\alpha \rfloor_0$, $\lfloor 1-\beta \rfloor_2$
x_1	$\lfloor \alpha \rfloor_1$, $\lfloor \gamma \rfloor_3$	$\lfloor 1-\alpha \rfloor_1$, $\lfloor 1-\gamma \rfloor_3$

Table 6: Lexicographic marginal probabilities from Table 3.

Definition 2 *Y is h-irrelevant to X given Z if*

$$\underline{E}[f(X)|A,B,z] = \underline{E}[f(X)|A,z] \quad \text{whenever } \{A,B,z\} \neq \emptyset.$$

X and Y are h-independent given Z when X is h-irrelevant to Y given Z and vice-versa.

We have:

Theorem 1 *H-independence satisfies Symmetry, Redundancy, Decomposition, and Weak Union.*

Proof. Symmetry holds by definition; Redundancy is trivial. From the assumed h-independence of X and (W,Y), we have: $\underline{E}[f(X)|A,B,z] = \underline{E}[f(X)|A,z]$, and $\underline{E}[g(Y)|A,B,z] = \underline{E}[g(Y)|B,z]$ (Decomposition). Weak Union follows from $\underline{E}[g(Y)|A,B,w,z] = \underline{E}[g(Y)|B,w,z]$, and then, using Decomposition, $\underline{E}[f(X)|A,w,z] = \underline{E}[f(X)|A,z] = \underline{E}[f(X)|A,B,w,z]$. ☐

Note that h-independence fails Contraction (Table 3).

In the next section we examine two other representations that are closely related to full conditional measures and full credal sets.

4 Lexicographic probabilities and sets of desirable gambles

Consider again Table 3. For this full distribution we have X and Y epistemic/h-/cs-/full/kr-/layer independent. One might argue that there is something strange about this "independence". For take a function $g(Y)$ such that $g(y_0) = -(1-\alpha)$ and $g(y_1) = \alpha$. This function has expected utility zero. But if $\beta < \alpha$ one might argue that g is better than the zero function; after all, if $\{w_1\}$ happens to be observed, then the expected value of g given $\{w_1\}$ is $\alpha - \beta$, and g should then be considered better than the zero function. And if $\gamma > \alpha$, then conditional on $\{w_1, x_1\}$ the zero function should be considered better than g. Hence conditioning on $\{x_1\}$ seems to change opinions about a function of Y.

One way to understand this example is to look at the marginal full probability for (X,Y), shown in Table 5. Note that when the full probability in Table 3 is marginalized over W, the content of layers L_2 and L_3 disappear: in Table 5 one sees neither β nor γ. Preferences about g that might depend on deeper layers can only be exposed by observing W. In a sense, the direct marginalization of Table 3 loses important information about the joint full probability. It would make more sense to say that the marginal probabilities obtained from Table 3 should be given by the overlapping layers in Table 6, so as to conclude that X and Y are *not* independent.

We are then moving into *lexicographic probabilities* that assign probability measures to various layers with possibly overlapping support. Due to the lack of space, we omit detailed background on lexicographic probabilities, and refer the reader to the work of Blume et al. [3] for all necessary definitions. We assume their axiomatization of the non-Archimedean preference relation $\succeq$, and use the fact that this preference relation can be represented by a sequence of probability measures over Ω; each one of these measures is a "layer" of the lexicographic probability. [3, Corollary 3.1]. Two functions $f_1(X)$ and $f_2(X)$ are compared with respect to a lexicographic rule in the sense that $f_1 \succeq f_2$ if and only if

$$\left[\sum_x f_1(x)P_i(x)\right]_{i=0}^{K} \geq_{\mathrm{L}} \left[\sum_x f_2(x)P_i(x)\right]_{i=0}^{K},$$

(for $a, b \in \Re^K$, $a \geq_{\mathrm{L}} b$ iff whenever $b_j > a_j$, there exists a $k < j$ such that $a_k > b_k$). These probabilities are unique only up to linear transformations, so there is some intrinsic non-uniqueness associated with lexicographic probabilities:

Example 2 Suppose that a binary variable Y is associated with two layers such that $P_0(y_0) = 1 - P_0(y_1) = \alpha$ and $P_1(y_0) = 1 - P_1(y_1) = \beta$. For fixed α, every $\beta \in [0, \alpha)$ yields identical preferences; likewise, every $\beta \in (\alpha, 1]$ yields identical preferences. So the specific value of β cannot be fixed by resorting to lexicographic preferences. $\square$

Conditional lexicographic probabilities given nonempty event H are obtained by conditioning every layer of the lexicographic probability on H, after discarding those layers that do not intersect H. These conditional probabilities encode the preferences $f_1(X)I_H \succeq f_2(X)I_H$ [3, Theorem 4.3], denoted by $[f_1(X) \succeq f_2(X)|H]$.

The close proximity between full probabilities and lexicographic probabilities is apparent. A full probability can be represented by a lexicographic probability with disjoint layers [22, 23]. And for any lexicographic probability, the function $P(A|B) = P_i(A|B)$, where P_i the the first measure such that $P_i(B) > 0$, is a full probability. However, as indicated by the discussion of marginalization concerning Tables 3, 5 and 6, full probabilities and lexicographic probabilities do not behave identically.

Now consider defining a concept of independence for lexicographic probabilities. We might try to define a "product" for lexicographic probabilities. Here difficulties abound due to non-uniqueness. First, probabilities in various layers can be modified so as to break factorization. Additionally, probability values are not tied to specific layer numbers. For instance, if we have a lexicographic probability with three overlapping layers, each with probability measures p_0, p_1 and p_2, we can generate an *equivalent* representation with four layers p_0, p_0, p_1 and p_2. Therefore a condition such as layer factorization seems rather fragile as we cannot control layer numbers just by looking at marginal lexicographic probabilities.

Indeed the difficulties with product lexicographic probabilities have already been discussed by several authors [3, 23, 24]. Solutions based on factorization of nonstandard measures have been advanced by these authors; the interpretation and the manipulation of such concepts do not seem easy, and we leave that to future work.

Hence we are led, in our study of lexicographic probabilities, to concepts of independence that rely on conditioning. Blume et al. [3] say that X and Y are independent when we have both

	w_0y_0	w_1y_0	w_0y_1	w_1y_1
x_0	$\lfloor\alpha\rfloor_0$	$\lfloor\beta\rfloor_2$	$\lfloor 1-\alpha\rfloor_0$	$\lfloor 1-\beta\rfloor_2$
x_1	$\lfloor\alpha\rfloor_1$	$\lfloor\beta\rfloor_3$, $\lfloor\gamma\rfloor_4$	$\lfloor 1-\alpha\rfloor_1$	$\lfloor 1-\beta\rfloor_3$, $\lfloor 1-\gamma\rfloor_4$

Table 7: Lexicographic distribution for W, X, Y, with distinct $\alpha \in (0,1)$, $\beta \in (0,1)$, $\gamma \in (0,1)$.

	y_0	y_1
x_0	$\lfloor\alpha\rfloor_0$, $\lfloor\beta\rfloor_2$	$\lfloor 1-\alpha\rfloor_0$, $\lfloor 1-\beta\rfloor_2$
x_1	$\lfloor\alpha\rfloor_1$, $\lfloor\beta\rfloor_3$	$\lfloor 1-\alpha\rfloor_1$, $\lfloor 1-\beta\rfloor_3$

$P(W, X|Y = y)$

	w_0	w_1
x_0	$\lfloor 1\rfloor_0$	$\lfloor 1\rfloor_2$
x_1	$\lfloor 1\rfloor_1$	$\lfloor 1\rfloor_3$

Table 8: Marginal (left) and conditional (right) lexicographic probabilities from Table 7.

$$[f_1(X) \succeq f_2(X)|y_1] \Leftrightarrow [f_1(X) \succeq f_2(X)|y_2],$$

$$[g_1(Y) \succeq g_2(Y)|x_1] \Leftrightarrow [g_1(Y) \succeq g_2(Y)|x_2]$$

whenever conditioning events are nonempty. Say that X and Y are independent given Z when the expressions above are satisfied conditional on any $\{z\}$ such that conditioning events are nonempty.

Even though Table 3 no longer fails Contraction if we use this concept of independence (because X and Y are no longer independent), consider Table 7. The distributions for (X, Y), for (X, W) given $\{y_0\}$, and for (X, W) given $\{y_1\}$ are shown in Table 8. Here X and Y are independent and X and W are independent given Y; yet X and (W, Y) are not independent. Contraction fails. The fourth layer "vanishes" when one marginalizes out W as preferences are decided already at the third layer. To understand this, consider Example 2: once α and β are fixed, every preference about Y is fixed, and there is no need to examine further layers.

Now suppose we have a *set* of lexicographic probabilities, where preference is given by unanimity amongst lexicographic comparisons [36]. Example 1 shows that Decomposition and Weak Union can fail for Blume et al.'s concept of independence (just consider each full probability a lexicographic probability, and take their convex hull if a convex set is desired).

We suggest that a more promising concept of independence for (sets of) lexicographic probabilities is obtained by symmetrizing the following concept: Y is irrelevant to X given Z when

$$[f_1(X) \succeq f_2(X)|A, B, z] \Leftrightarrow [f_1(X) \succeq f_2(X)|A, z],$$

for all functions, whenever conditioning events are nonempty. And X and Y are independent given Z when Y is irrelevant to X given Z and vice-versa.

This concept of independence satisfies Symmetry, Redundancy, Decomposition and Weak Union; Contraction fails (Table 7). Redundancy obtains because

$$\begin{aligned}[f_1(X) \succeq f_2(X)|A, B, x] \quad &\Leftrightarrow \quad f_1(x) \geq f_2(x) \\ &\Leftrightarrow \quad [f_1(X) \succeq f_2(X)|B, x].\end{aligned}$$

Decomposition holds because any event B belongs to the algebra generated by (W, Y), and any function $g(Y)$ is also a function of (W, Y) (hence independence of X and (W, Y) given Z implies independence of X and Y given Z). Weak Union holds because, assuming X and (W, Y) independent given Z, we have

$$[g_1(Y) \succeq g_2(Y)|A, B, w, z] \Leftrightarrow [g_1(Y) \succeq g_2(Y)|B, w, z],$$

and, using Decomposition,

$$\begin{aligned}[f_1(X) \succeq f_2(X)|A, w, z] \quad &\Leftrightarrow \quad [f_1(X) \succeq f_2(X)|A, z] \\ &\Leftrightarrow \quad [f_1(X) \succeq f_2(X)|A, B, w, z].\end{aligned}$$

Sets of lexicographic probabilities are equivalent, from the point of view of preference representations, to *sets of desirable gambles*, a representation that has received considerable attention [6, 17, 18, 31, 43]. Indeed the derivation of lexicographic representations for sets of desirable gambles appears already in the work of Seidenfeld et al. [36], who show that a partially ordered set of preferences (that encodes a set of desirable gambles) can be represented by a set of complete orderings, each one of which can be represented by a lexicographic probability (either using results by Kee [25] or the more direct results by Blume et al. [3]). In recent work, Couso and Moral [6] have studied the representation of sets of desirable gambles through lexicographic probabilities.

A set of desirable gambles $\mathbb{D}$ is a set of variables not containing the zero function and containing all nonnegative variables that are different from zero, and such that $\lambda X \in \mathbb{D}$ if $X \in \mathbb{D}$ and $\lambda > 0$, and $X + Y \in \mathbb{D}$ if $X, Y \in \mathbb{D}$ [17, Definition 1]. The set of desirable gambles conditional on event A, denoted by $[\mathbb{D}|A]$, contains all desirable gambles X such that $XI_A = X$, where I_A is the indicator function of A [18, Section 3.2]. Following notation by Moral [31], denote by $\mathbb{D}^{\downarrow X}$ the set of desirable gambles that are functions of X (that is, $\mathbb{D}^{\downarrow X}$ is the "marginal" set of gambles with respect to X). A natural concept of independence for sets of desirable gambles is [17, Definition 3]: Y is irrelevant to X given Z if

$$[\mathbb{D}|y, z]^{\downarrow X} = [\mathbb{D}|z]^{\downarrow X} \text{ whenever } \{y, z\} \neq \emptyset.$$

And then: X and Y are independent given Z if X is irrelevant to Y given Z and vice-versa. (Note that there are other concepts of independence for sets of desirable gambles in the literature [31].)

Mimicking our proposal for (sets of) lexicographic probabilities, consider the following definition of independence for sets of desirable gambles: Y is irrelevant to X if

$$[\mathbb{D}|A, B, z]^{\downarrow X} = [\mathbb{D}|A, z]^{\downarrow X} \text{ whenever } \{A, B, z\} \neq \emptyset.$$

And then define independence of X and Y given Z by symmetrizing this concept of irrelevance.

5 Conclusion

This paper has studied concepts of independence for sets of full probabilities, and for their close relatives, sets of lexicographic probabilities, and sets of desirable gambles. We have tried to offer a commented and organized review of the literature in Section 2. We have then analyzed a large number of concepts of independence in Sections 3 and 4.

At this point the only concept of independence for full credal sets that satisfy Symmetry, Redundancy, Decomposition, Weak Union and Contraction is elementwise layer independence. The concepts of confirmational and epistemic independence seem particularly weak when applied to full credal sets. The concept of h-independence fares considerably better but still fails Contraction. The extent to which one can adopt concepts that fail various graphoid properties is yet to be fully analyzed.

Concerning lexicographic probabilities: they do add flexibility, but they introduce significant complexity in dealing with non-uniqueness and marginalization. Sets of desirable gambles also require some care in dealing with marginalization. The new concepts of independence suggested here for sets of lexicographic probabilities and sets of desirable gambles should be helpful in future work.

Acknowledgements

Thanks to reviewers, who suggested many improvements to the original submission.

The author is partially supported by CNPq.

References

[1] Pierpaolo Battigalli. Strategic independence and perfect Bayesian equilibria. *Journal of Economic Theory*, 70:201–234, 1996.

[2] Pierpaolo Battigalli and Pietro Veronesi. A note on stochastic independence without Savage-null

events. *Journal of Economic Theory*, 70(1):235–248, 1996.

[3] Lawrence Blume, Adam Brandenburger, and Eddie Dekel. Lexicographic probabilities and choice under uncertainty. *Econometrica*, 58(1):61–79, January 1991.

[4] Giulianella Coletti and Romano Scozzafava. Stochastic independence in a coherent setting. *Annals of Mathematics and Artificial Intelligence*, 35:151-176, 2002.

[5] Giulianella Coletti and Romano Scozzafava. *Probabilistic Logic in a Coherent Setting*. Trends in logic, 15. Kluwer, Dordrecht, 2002.

[6] Inés Couso and Serafín Moral. Sets of desirable gambles: Conditioning, representation, and precise probabilities. *International Journal of Approximate Reasoning*, 52:1034–1055, 2011.

[7] Fabio Gagliardi Cozman. Irrelevance and independence axioms in quasi-Bayesian theory. In Anthony Hunter and Simon Parsons, editors, *European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU)*, pages 128–136. Springer, London, England, 1999.

[8] Fabio Gagliardi Cozman. Computing lower expectations with Kuznetsov's independence condition. In *Proceedings of the Third International Symposium on Imprecise Probabilities and Their Applications*, pages 177–187, Lugano, Switzerland, 2003. Carleton Scientific.

[9] Fabio Gagliardi Cozman. Sets of probability distributions, independence, and convexity. *Synthese*, 186(2):577–600, 2012.

[10] Fabio Gagliardi Cozman. Independence for Full Conditional Probabilities: Structure, Nonuniqueness, Factorization, and Bayesian Networks. Technical Report Decision Making Lab 2013-001, São Paulo, Brazil, 2013.

[11] Fabio Gagliardi Cozman and Teddy Seidenfeld. Independence for full conditional measures and their graphoid properties. In Benedikt Lowe, Eric Pacuit, and Jan-Willem Romeijn, editors, *Reasoning about Probabilities and Probabilistic Reasoning*, volume 16 of *Foundations of the Formal Sciences VI*, pages 1–29. College Publications, London, 2009.

[12] Fabio Gagliardi Cozman and Peter Walley. Graphoid properties of epistemic irrelevance and independence. *Annals of Mathematics and Artificial Intelligence*, 45:173–195, 2005.

[13] Ákos Császár. Sur la structure des espaces de probabilité conditionnelle. *Acta Mathematica Academiae Scientiarum Hungarica*, 6(3-4):337–361, 1955.

[14] A. Philip Dawid. Conditional independence in statistical theory (with discussion). *Journal of the Royal Statistical Society B*, 41:1–31, 1979.

[15] L. de Campos and Serafín Moral. Independence concepts for convex sets of probabilities. In Phillippe Besnard and Steve Hanks, editors, *XI Conference on Uncertainty in Artificial Intelligence*, pages 108–115, San Francisco, California, United States, 1995. Morgan Kaufmann.

[16] Jasper De Bock and Gert De Cooman. Imprecise Bernoulli processes. In *Communications in Computer and Information Science*, volume 299, pages 400–409. Springer, 2012.

[17] Gert de Cooman and Enrique Miranda. Irrelevant and independent natural extension for sets of desirable gambles. *Journal of Artificial Intelligence Research*, 45:601–640, 2012.

[18] Gert de Cooman and Erik Quaeghebeur. Exchangeability and sets of desirable gambles. *International Journal of Approximate Reasoning*, 53:363–395, 2012.

[19] Bruno de Finetti. *Theory of Probability, vol. 1-2.* Wiley, New York, 1974.

[20] Lester E. Dubins. Finitely additive conditional probability, conglomerability and disintegrations. *Annals of Statistics*, 3(1):89–99, 1975.

[21] F. J. Giron and S. Rios. Quasi-Bayesian behaviour: A more realistic approach to decision making? In J. M. Bernardo, J. H. DeGroot, D. V. Lindley, and A. F. M. Smith, editors, *Bayesian Statistics*, pages 17–38. University Press, Valencia, Spain, 1980.

[22] Joseph Y. Halpern. Lexicographic probability, conditional probability, and nonstandard probability. *Games and Economic Behavior*, 68:155–179, 2010.

[23] Peter J. Hammond. Elementary non-Archimedean representations of probability for decision theory and games. In P. Humphreys, editor, *Patrick Suppes: Scientific Philosopher; Volume 1*, pages 25–59. Kluwer, Dordrecht, The Netherlands, 1994.

[24] Peter J. Hammond. Non-Archimedean subjective probabilities in decision theory and games. Technical Report Working Paper No. 97-038, Stanford University Department of Economics, 1997.

[25] V. L. Klee Jr. The structure of semispaces. *Mathematica Scandinavia*, 4:54-64, 1956.

[26] Elon Kohlberg and Philip J. Reny. Independence on relative probability spaces and consistent assessments in game trees. *Journal of Economic Theory*, 75:280–313, 1997.

[27] Peter Krauss. Representation of conditional probability measures on Boolean algebras. *Acta Mathematica Academiae Scientiarum Hungaricae*, 19(3-4):229–241, 1968.

[28] V. P. Kuznetsov. *Interval Statistical Methods*. Radio i Svyaz Publ., (in Russian), 1991.

[29] Isaac Levi. *The Enterprise of Knowledge*. MIT Press, Cambridge, Massachusetts, 1980.

[30] V. McGee. Learning the impossible. In E. Bells and B. Skyrms, editors, *Probability and Conditionals*, pages 179–199. Cambridge University Press, 1994.

[31] Serafin Moral. Epistemic irrelevance on sets of desirable gambles. *Annals of Mathematics and Artificial Intelligence*, 45(1-2):197–214, October 2005.

[32] Roger B. Myerson. *Game Theory: Analysis of Conflict*. Harvard University Press, Cambridge, MA, 1991.

[33] Roger B. Myerson. Multistage games with communication. *Econometrica*, 54(2):323–358, 1986.

[34] Judea Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, San Mateo, California, 1988.

[35] Karl Raimund Popper. *The Logic of Scientific Discovery*. Hutchinson, London, 1975.

[36] Teddy Seidenfeld, Mark J. Schervish, and Joseph B. Kadane. Decisions without ordering. In W. Sieg, editor, *Acting and Reflecting*, pages 143–170. Kluwer Academic Publishers, 1990.

[37] Teddy Seidenfeld, Mark J. Schervish, and Joseph Kadane. Coherent choice functions under uncertainty. In *International Symposium on Imprecise Probability: Theories and Applications*. Prague, Czech Republic, 2007.

[38] Bas Cornelis van Fraassen. Fine-grained opinion, probability, and the logic of full belief. *Journal of Philosophical Logic*, 24:349–377, 1995.

[39] Barbara Vantaggi. Conditional independence in a coherent finite setting. *Annals of Mathematics and Artificial Intelligence*, 32(1-4):287–313, 2001.

[40] Barbara Vantaggi. Conditional independence structures and graphical models. *International Journal of Uncertainty, Fuzziness and Knoweledge-Based Systems*, 11(5):545-571, 2003.

[41] Peter Walley. Coherent lower (and upper) probabilities. Technical Report Statistics Report 23, University of Warwick, Coventry, 1981.

[42] Peter Walley. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London, 1991.

[43] Peter Walley. Towards a unified theory of imprecise probability. *International Journal of Approximate Reasoning*, 24:125–148, 2000.

[44] Kurt Weichselberger. The theory of interval-probability as a unifying concept for uncertainty. *International Journal of Approximate Reasoning*, 24(2-3):149–170, 2000.

[45] Kurt Weichselberger, T. Augustin (assistant), and A. Wallner (assistant). *Elementare Grundbegriffe einer allgemeineren Wahrscheinlichkeitsrechnung I: Intervallwahrscheinlichkeit als umfassendes Konzept*. Physica-Verlag Heidelberg, 2001.