Conditioning, Conditional Independence and Irrelevance in Evidence Theory

Jiřina Vejnarová
Institute of Information Theory and Automation
Academy of Sciences of the Czech Republic
vejnar@utia.cas.cz

Abstract

The goal of the paper is to reveal the relationships between recently introduced concept of conditional independence in evidence theory and those (dependent on the choice of conditioning rule) of conditional irrelevance.

Keywords. Evidence theory, multidimensional models, conditioning rules, conditional independence, conditional irrelevance.

1 Introduction

When applying models of artificial intelligence to any practical problem one must cope with two basic problems: uncertainty and multidimensionality. The most widely used models managing these issues are, at present, so-called *probabilistic graphical Markov models*.

The problem of multidimensionality is solved in these models with the help of the notion of conditional independence, which enables factorization of a multidimensional probability distribution into small parts, usually marginal or conditional low-dimensional distributions (e.g. in *Bayesian networks*), or generally into low-dimensional factors (e.g. in *decomposable models*). Such a factorization not only decreases the storage requirements for representation of a multidimensional distribution but it usually also induces efficient computational procedures allowing inference from these models.

It is easy to realize that if we need efficient methods for representation of probability distributions (requiring an exponential number of parameters), the greater is the need of an efficient tool for representation of belief functions, which cannot be represented by a distribution (but only by a set function), and therefore the space requirements for its representation are superexponential. To solve this problem, in [9, 15] we proposed a new concept of conditional independence in evidence theory, proved its formal properties and showed [16] in which sense it is superior to the previous one [3].

However, another problem appears when one tries to construct an evidential counterpart of Bayesian network: problem of conditioning, which is not sufficiently solved in evidence theory. There exist many conditioning rules [6], but is any of them compatible with our conditional independence concept? In other words, if one is interested in Bayesian-networks-like evidential models, he/she will need rather the concept of conditional irrelevance. Therefore, it is also necessary to find the relationship between conditional independence and irrelevance. It is not necessary for Bayesian networks, as in (precise) probability framework the difference between conditional independence and irrelevance is only subtle.

The contribution is organized as follows. After a short overview of necessary terminology and notation (Section 2), in Section 3 we recall two conditioning rules (suggested for conditioning of events) and introduce their generalizations for variables. In Section 4 the above-mentioned concept of conditional independence is recalled and a new concept of (conditional) irrelevance is presented. In Section 5 the relationship between (conditional) independence and (conditional) irrelevance is studied.

2 Basic Concepts

In this section we will briefly recall basic concepts from evidence theory [11] concerning sets, set functions and marginalization.

2.1 Set projections and extensions

For an index set $N = \{1, 2, \ldots, n\}$ let $\{X_i\}_{i\in N}$ be a system of variables, each X_i having its values in a finite set $\mathbf{X}_i$. In this paper we will deal with *multidi-*

mensional frame of discernment

$$\mathbf{X}_N = \mathbf{X}_1 \times \mathbf{X}_2 \times \ldots \times \mathbf{X}_n,$$

and its *subframes* (for $K \subseteq N$)

$$\mathbf{X}_K = \times_{i \in K} \mathbf{X}_i.$$

When dealing with groups of variables on these subframes, X_K will denote a group of variables $\{X_i\}_{i\in K}$ throughout the paper.

A *projection* of $x = (x_1, x_2, \ldots, x_n) \in \mathbf{X}_N$ into $\mathbf{X}_K$ will be denoted $x^{\downarrow K}$, i.e. for $K = \{i_1, i_2, \ldots, i_k\}$

$$x^{\downarrow K} = (x_{i_1}, x_{i_2}, \ldots, x_{i_k}) \in \mathbf{X}_K.$$

Analogously, for $M \subset K \subseteq N$ and $A \subset \mathbf{X}_K$, $A^{\downarrow M}$ will denote a *projection* of A into $\mathbf{X}_M$:[1]

$$A^{\downarrow M} = \{y \in \mathbf{X}_M \mid \exists x \in A : y = x^{\downarrow M}\}.$$

In addition to the projection, in this text we will also need an opposite operation usually called a cylindrical extension. The *cylindrical extension* of $A \subset \mathbf{X}_K$ to $\mathbf{X}_L$ $(K \subset L)$ is the set

$$A^{\uparrow L} = \{x \in \mathbf{X}_L : x^{\downarrow K} \in A\}.$$

Clearly

$$A^{\uparrow L} = A \times \mathbf{X}_{L\setminus K}.$$

A more complicated case is to make common extension of two sets, which will be called a join. By a *join*[2] of two sets $A \subseteq \mathbf{X}_K$ and $B \subseteq \mathbf{X}_L$ $(K, L \subseteq N)$ we will understand a set

$$A \bowtie B = \{x \in \mathbf{X}_{K\cup L} : x^{\downarrow K} \in A \;\;\&\;\; x^{\downarrow L} \in B\}.$$

Let us note that for any $C \subseteq \mathbf{X}_{K\cup L}$ naturally $C \subseteq C^{\downarrow K} \bowtie C^{\downarrow L}$, but generally $C \neq C^{\downarrow K} \bowtie C^{\downarrow L}$.

Let us also note that if K and L are disjoint, then the join of A and B is just their Cartesian product $A \bowtie B = A \times B$, if $K = L$ then $A \bowtie B = A \cap B$. If $K \cap L \neq \emptyset$ and $A^{\downarrow K\cap L} \cap B^{\downarrow K\cap L} = \emptyset$ then also $A \bowtie B = \emptyset$. Generally,

$$A \bowtie B = (A \times \mathbf{X}_{L\setminus K}) \cap (B \times \mathbf{X}_{K\setminus L}),$$

i.e. a join of two sets is the intersection of their cylindrical extensions.

[1]Let us remark that we do not exclude situations when $M = \emptyset$. In this case $A^{\downarrow\emptyset} = \emptyset$.

[2]This term and notation are taken from the theory of relational databases [1].

2.2 Set functions

In evidence theory [11] (or Dempster-Shafer theory) two measures are used to model the uncertainty: belief and plausibility measures. Both of them can be defined with the help of another set function called a *basic (probability* or *belief) assignment* m on $\mathbf{X}_N$, i.e. ,

$$m : \mathcal{P}(\mathbf{X}_N) \longrightarrow [0, 1],$$

where $\mathcal{P}(\mathbf{X}_N)$ is power set of $\mathbf{X}_N$ and

$$\sum_{A\subseteq \mathbf{X}_N} m(A) = 1.$$

Furthermore, we assume that $m(\emptyset) = 0$.

A set $A \in \mathcal{P}(\mathbf{X}_N)$ is a *focal element* if $m(A) > 0$. Let $\mathcal{F}$ denote the set of all focal elements, a focal element $A \in \mathcal{F}$ is called an $m-$*atom* if for any $B \subseteq A$ either $B = A$ or $B \notin \mathcal{F}$. In other words, $m-$*atom* is a setwise-minimal focal element.

Let us note that atomicity of a focal element is not closed with respect to either marginalization or extension.

Belief and *plausibility measures* are defined for any $A \subseteq \mathbf{X}_N$ by the equalities

$$Bel(A) = \sum_{B\subseteq A} m(B). \quad (1)$$

$$Pl(A) = \sum_{B\cap A\neq\emptyset} m(B), \quad (2)$$

respectively.

It is well-known (and evident from these formulae) that for any $A \in \mathcal{P}(\mathbf{X}_N)$

$$Bel(A) \leq Pl(A), \quad (3)$$

$$Pl(A) = 1 - Bel(A^C), \quad (4)$$

where A^C is the set complement of $A \in \mathcal{P}(\mathbf{X}_N)$. Furthermore, basic assignment can be computed from belief function via Möbius inversion:

$$m(A) = \sum_{B\subseteq A} (-1)^{|A\setminus B|} Bel(B), \quad (5)$$

i.e. any of these three functions is sufficient to define values of the remaining two.

2.3 Marginalization

For a basic assignment m on $\mathbf{X}_K$ and $M \subset K$, a *marginal basic assignment* of m on $\mathbf{X}_M$ is defined (for each $A \subseteq \mathbf{X}_M$):

$$m^{\downarrow M}(A) = \sum_{\substack{B\subseteq \mathbf{X}_K \\ B^{\downarrow M}=A}} m(B). \quad (6)$$

Analogously we will denote by $Bel^{\downarrow M}$ and $Pl^{\downarrow M}$ marginal belief and plausibility measures on $\mathbf{X}_M$, respectively.

The following simple lemma concerning marginal beliefs and plausibilities will be used in the next section.

Lemma 1 *Let m be a basic assignment on $\mathbf{X}_N$, Bel and Pl corresponding beliefs and plausibilities and $K \subset N$. Then for any $A \subset \mathbf{X}_K$*

$$\begin{aligned} Bel^{\downarrow K}(A) &= Bel(A^{\uparrow N}), \qquad (7)\\ Pl^{\downarrow K}(A) &= Pl(A^{\uparrow N}). \qquad (8) \end{aligned}$$

Proof. Using (1) and (6) one obtains

$$\begin{aligned} Bel^{\downarrow K}(A) &= \sum_{\substack{B\subseteq \mathbf{X}_K\\ B\subseteq A}} m^{\downarrow K}(B)\\ &= \sum_{\substack{B\subseteq \mathbf{X}_K\\ B\subseteq A}} \sum_{\substack{C\subseteq \mathbf{X}_N\\ C^{\downarrow K}=B}} m(C)\\ &= \sum_{\substack{C\subseteq \mathbf{X}_N\\ C^{\downarrow K}\subseteq A}} m(C)\\ &= \sum_{\substack{C\subseteq \mathbf{X}_N\\ C\subseteq A^{\uparrow N}}} m(C)\\ &= Bel(A^{\uparrow N}), \end{aligned}$$

where we used the fact that $C^{\downarrow K} \subseteq A$ if and only if $C \subseteq A^{\uparrow N}$ for any $C \subseteq \mathbf{X}_N$ and $A \subseteq \mathbf{X}_K$.

Similarly, using (2), (6) and the fact that $D^{\downarrow K} \subseteq \mathbf{X}_K, D^{\downarrow K} \cap B \neq \emptyset$ if and only if $D \subseteq \mathbf{X}_N, D \cap B^{\uparrow N} \neq \emptyset$

$$\begin{aligned} Pl^{\downarrow K}(B) &= \sum_{\substack{C\subseteq \mathbf{X}_K\\ C\cap B\neq\emptyset}} m^{\downarrow K}(C)\\ &= \sum_{\substack{C\subseteq \mathbf{X}_K\\ C\cap B\neq\emptyset}} \sum_{\substack{D\subseteq \mathbf{X}_N\\ D^{\downarrow K}=C}} m(D)\\ &= \sum_{\substack{D\subseteq \mathbf{X}_N\\ D\cap B^{\uparrow L}\neq\emptyset}} m(D)\\ &= Pl(B^{\uparrow N}), \end{aligned}$$

as desired. □

3 Conditioning

Conditioning belongs to the most important topics of any theory dealing with uncertainty. From the viewpoint of construction of Bayesian-network-like multidimensional models it seems to be inevitable.

3.1 Conditioning of Events

In evidence theory the "classical" conditioning rule is so-called *Dempster's rule of conditioning* defined for any $\emptyset \neq A \subseteq \mathbf{X}_N$ and $B \subseteq \mathbf{X}_N$ such that $Pl(B) > 0$ by the formula

$$m(A|B) = \frac{\sum_{C\subseteq \mathbf{X}_N : C\cap B = A} m(C)}{Pl(B)} \qquad (9)$$

and $m(\emptyset|B) = 0$.

Let us note that formula (9) is special case of Dempster's rule of combination, when combining basic assignment m with another m_B such that $m_B(B) = 1$.

From this formula one can immediately obtain:

$$\begin{aligned} Bel(A|B) &= \frac{Bel(A\cup B^C) - Bel(B^C)}{1 - Bel(B^C)},\\ Pl(A|B) &= \frac{Pl(A\cap B)}{Pl(B)}. \qquad (10) \end{aligned}$$

This is not the only possibility how to make conditioning, another — in a way symmetric — conditioning rule is the following one called *focusing* defined for any $\emptyset \neq A \subseteq \mathbf{X}_N$ and $B \subseteq \mathbf{X}_N$ such that $Bel(B) > 0$ by the formula

$$m(A||B) = \begin{cases} \dfrac{m(A)}{Bel(B)} & \text{if} \quad A \subseteq B,\\ 0 & \text{otherwise.} \end{cases}$$

From the following two equalities one can see, in which sense are these two conditioning rules symmetric:

$$\begin{aligned} Bel(A||B) &= \frac{Bel(A\cap B)}{Bel(B)}, \qquad (11)\\ Pl(A||B) &= \frac{Pl(A\cup B^C) - Pl(B^C)}{1 - Pl(B^C)}. \end{aligned}$$

These rules are based on different philosophy. Focusing assigns positive values only to those elements which are subsets of B, while Dempster's rule of conditioning to those which have nonempty intersection with it.

It is evident, that focusing is applicable in less cases than Dempster's rule, because of relation (3), hence from this point of view the latter seems to be more advantageous.

On the other hand, from the computational viewpoint the latter is more suitable, as it produces less focal elements (and in any of them a bigger "mass" is contained; cf. also Example 1). Due to this fact it may seem that focusing produces bigger intervals than Dempster's rule (and it is very often true), but it is not generally satisfied, as can be seen again from Example 1.

Formulae (10) and (11) are, in a way, evidential counterparts of conditioning in probabilistic framework. Let us note that seemingly "natural" way of conditioning

$$m(A|_pB) = \frac{m(A \cap B)}{m(B)} \tag{12}$$

is not possible, since $m(A|_pB)$ need not be a basic assignment, as can be seen from the following simple example. It is caused by a simple fact that m, in contrary to Bel and Pl is not monotonous with respect to set inclusion.

Example 1 Let $\mathbf{X} = \{a, b, c\}$ and m on $\mathbf{X}$ be defined as follows:

$$\begin{aligned} m(\{a\}) = m(\{b\}) = m(\{c\}) &= \frac{1}{4}, \\ m(\{a, b\}) = m(\mathbf{X}) &= \frac{1}{8}. \end{aligned}$$

Using (12) one would obtain

$$m(\{a\}|_p\{a, b\}) = m(\{b\}|_p\{a, b\}) = 2,$$

which is out of the framework of evidence theory.

Let us use this example also for demonstrating the difference between Dempster's rule of conditioning and focusing. For this purpose let us compute

$$Bel(\{b, c\} = \frac{1}{2} \quad \text{and} \quad Pl(\{b, c\} = \frac{3}{4}.$$

Then we have

$$\begin{aligned} m(\{b\}|\{b, c\}) = \frac{m(\{b\}) + m(\{a, b\})}{Pl(\{b, c\})} &= \frac{1}{2}, \\ m(\{c\}|\{b, c\}) = \frac{m(\{c\})}{Pl(\{b, c\})} &= \frac{1}{3}, \\ m(\{b, c\}|\{b, c\}) = \frac{m(\mathbf{X})}{Pl(\{b, c\})} &= \frac{1}{6}, \end{aligned}$$

as $\{a, b\} \cap \{b, c\} = \{b\}$ and $\mathbf{X} \cap \{b, c\} = \{b, c\}$, while

$$\begin{aligned} m(\{b\}||\{b, c\}) = \frac{m(\{b\})}{Bel(\{b, c\})} &= \frac{1}{2}, \\ m(\{c\}||\{b, c\}) = \frac{m(\{c\})}{Bel(\{b, c\})} &= \frac{1}{2}, \end{aligned}$$

as $\{b\}$ and $\{c\}$ are the only subsets of $\{b, c\}$. ♢

Nevertheless, rather than in conditional beliefs and plausibilities of events we are interested in conditioning by variables. This problem will be in the center of our attention in the next subsection.

3.2 Conditional Variables

Definition 1 Let X_K and X_L ($K \cap L = \emptyset$) be two groups of variables with values in $\mathbf{X}_K$ and $\mathbf{X}_L$, respectively. Then the *conditional basic assignment according to Dempster's conditioning rule* of X_K given $X_L \in B \subseteq \mathbf{X}_L$ (for B such that $Pl(B) > 0$) is defined as follows:

$$m_{X_K|X_L}(A|B) = \frac{\sum_{C \subseteq \mathbf{X}_{K \cup L} : (C \cap B^{\uparrow K \cup L})^{\downarrow K} = A} m(C)}{Pl(B)} \tag{13}$$

for $A \neq \emptyset$ and $m_{K|L}(\emptyset|B) = 0$. Similarly, the *conditional basic assignment according to focusing* of X_K given $X_L \in B \subseteq \mathbf{X}_L$ (for B such that $Bel(B) > 0$) is defined by the equality

$$m_{X_K||X_L}(A||B) = \frac{\sum_{C \subseteq \mathbf{X}_{K \cup L} : C \subseteq B^{\uparrow K \cup L} \& C^{\downarrow K} = A} m(C)}{Bel(B)} \tag{14}$$

for any $A \neq \emptyset$ and $m_{K||L}(\emptyset||B) = 0$.

Now, let us prove that the definition is correct.

Theorem 1 *Set functions $m_{X_K|X_L}$ and $m_{X_K||X_L}$ defined for any fixed $B \subseteq \mathbf{X}_L$, such that $Pl(B) > 0$ and $Bel(B) > 0$, respectively, by Definition 1 are basic assignments on $\mathbf{X}_K$.*

Proof.

(i) Let $B \subseteq \mathbf{X}_L$ be such that $Pl(B) > 0$. As nonnegativity of $m_{X_K|X_L}(A|B)$ for any $A \subseteq \mathbf{X}_K$ and the fact that $m_{X_K|X_L}(\emptyset|B) = 0$ follow directly from the definition, to prove that $m_{X_K|X_L}$ is a basic assignment it is enough to show that

$$\sum_{A \subseteq \mathbf{X}_K} m_{X_K|X_L}(A|B) = 1.$$

To check it, let us sum the values of the numerators in (13)

$$\begin{aligned} \sum_{A \subseteq \mathbf{X}_K} \sum_{\substack{C \subseteq \mathbf{X}_{K \cup L} \\ (C \cap B^{\uparrow K \cup L})^{\downarrow K} = A}} m(C) &= \sum_{\substack{A \subseteq \mathbf{X}_K \\ A \neq \emptyset}} \sum_{\substack{C \subseteq \mathbf{X}_{K \cup L} \\ (C \cap B^{\uparrow K \cup L})^{\downarrow K} = A}} m(C) \\ &= \sum_{\substack{C \subseteq \mathbf{X}_{K \cup L} \\ C \cap B^{\uparrow L} \neq \emptyset}} m(C) \\ &= Pl(B^{\uparrow L}). \end{aligned}$$

To finish the proof it is enough to realize that $Pl(B^{\uparrow K\cup L}) = Pl^{\downarrow L}(B)$ for any $B \subseteq \mathbf{X}_L$ (by (8) of Lemma 1).

(ii) Analogously we will show that $m_{X||Y}$ is defined correctly. Let $B \subseteq \mathbf{X}_L$ be such that $Bel(B) > 0$. To prove that $m_{X||Y}$ is a basic assignment, it is again enough to check that

$$\sum_{A\subseteq \mathbf{X}_K} m_{X_K||X_L}(A||B) = 1.$$

To do so, let us compute

$$\begin{aligned}\sum_{A\subseteq \mathbf{X}_K} \sum_{\substack{C\subseteq \mathbf{X}_{K\cup L}:\\ C\subseteq B^{\uparrow K\cup L}: C^{\downarrow K}=A}} m(C) &= \sum_{\substack{C\subseteq \mathbf{X}_{K\cup L}:\\ C\subseteq B^{\uparrow L}A}} m(C) \\ &= Bel(B^{\uparrow L}).\end{aligned}$$

The rest of the proof, i.e. validity of $Bel(B^{\uparrow K\cup L}) = Bel^{\downarrow L}(B)$ follows directly from (7) of Lemma 1. □

4 Conditional Independence and Irrelevance

4.1 Conditional Independence and Irrelevance in Probability Theory

Independence and irrelevance need not be (and usually are not) distinguished in the probabilistic framework, as they are almost equivalent to each other.

Supposing X_K, X_L and X_M are groups of random variables with a joint probability distribution P we say that X_K is *conditionally independent* of X_L given X_M with respect to P if the equality

$$\begin{aligned}P(x_K, x_L, x_M) \cdot P^{\downarrow M}(x_M) \\ = P^{\downarrow K\cup M}(x_K, x_M) \cdot P^{\downarrow L\cup M}(x_L, x_M)\end{aligned}$$

(where $P_{X_K X_M}, P_{X_L X_M}, P_{X_M}$ denote corresponding marginal distributions) holds for every value (x_K, x_L, x_M) of the variables X_K, X_L, X_M. It means that in every situation when the value of X_M is known the values of X_K and X_L are completely unrelated (from the stochastic point of view).

There exist several equivalent definitions of stochastic conditional independence, e.g.

$$P_{X_K|X_L X_M}(x_K|x_L, x_M) = P_{X_K|X_M}(x_K|x_M),$$

but this definition may be used only in the situation when $P^{\downarrow L\cup M}(x_L, x_M)$ is positive.

Similarly, in possibilistic framework adopting De Cooman's measure-theoretical approach [7] (particularly his notion of almost everywhere equality) we proved that analogous definitions are equivalent (for more details see [13]).

4.2 Independence

When constructing graphical models in any framework, (conditional) independence concept plays an important role. In evidence theory the most common notion of independence is that of random set independence [5].

It has already been proven [14] that it is also the only sensible one, as e.g. application of strong independence to two bodies of evidence may generally lead to a model which is beyond the framework of evidence theory. Epistemic independence and irrelevance were not taken into consideration, as none of them seem to be a suitable tool for factorization of multidimensional models. Furthermore, they require conditioning, so their application is also problematic from this point of view.

Definition 2 Let m be a basic assignment on $\mathbf{X}_N$ and $K, L \subset N$ be disjoint. We say that groups of variables X_K and X_L are *independent with respect to basic assignment* m (in notation $K \perp\!\!\!\perp L\ [m]$) if

$$m^{\downarrow K\cup L}(A) = m^{\downarrow K}(A^{\downarrow K}) \cdot m^{\downarrow L}(A^{\downarrow L})$$

for all $A \subseteq \mathbf{X}_{K\cup L}$ for which $A = A^{\downarrow K} \times A^{\downarrow L}$, and $m(A) = 0$ otherwise.

This notion can be generalized in various ways [3, 12, 15]; the concept of conditional non-interactivity from [3], based on conjunction combination rule, is used for construction of directed evidential networks in [4]. In this paper we will use the concept introduced in [9, 15], as we consider it more suitable (the arguments can be found in [15]).

Definition 3 Let m be a basic assignment on $\mathbf{X}_N$ and $K, L, M \subset N$ be disjoint, $K \neq \emptyset \neq L$. We say that groups of variables X_K and X_L are *conditionally independent given* X_M *with respect to* m (and denote it by $K \perp\!\!\!\perp L|M\ [m]$), if the equality

$$\begin{aligned}m^{\downarrow K\cup L\cup M}(A) \cdot m^{\downarrow M}(A^{\downarrow M}) \\ = m^{\downarrow K\cup M}(A^{\downarrow K\cup M}) \cdot m^{\downarrow L\cup M}(A^{\downarrow L\cup M})\end{aligned}$$

holds for any $A \subseteq \mathbf{X}_{K\cup L\cup M}$ such that $A = A^{\downarrow K\cup M} \bowtie A^{\downarrow L\cup M}$, and $m(A) = 0$ otherwise.

It has been proven in [15] that this conditional independence concept satisfies so-called semi-graphoid properties taken as reasonable to be valid for any conditional independence concept (see e.g. [10]) and it has been shown in which sense this conditional independence concept is superior to previously introduced ones [3, 12].

4.3 Irrelevance

Irrelevance is usually considered to be a weaker notion than independence (see e.g. [5]). It expresses the fact that a new piece of evidence concerning one variable cannot influence the evidence concerning the other variable, in other words is irrelevant to it. More formally: group of variables X_L is *irrelevant* to X_K ($K \cap L = \emptyset$) if for any $B \subseteq \mathbf{X}_L$ such that $Pl(B) > 0$

$$m_{X_K|X_L}(A|B) = m(A) \qquad (15)$$

for any $A \subseteq \mathbf{X}_K$.[3]

It follows from the definition of irrelevance that it need not be a symmetric relation. Its symmetrized version is sometimes taken as a definition of independence. Let us note, that in the framework of evidence theory even in cases when the relation is symmetric, it does not imply independence, as can be seen from Examples 2 and 3.

Generalization of this notion to conditional irrelevance may be done as follows. Group of variables X_L is *conditionally irrelevant* to X_K given X_M (K, L, M disjoint, $K \neq \emptyset \neq L$) if for any $B \subseteq \mathbf{X}_L$ and $C \subseteq \mathbf{X}_M$ such that $Pl(B \times C) > 0$

$$m_{X_K|X_L X_M}(A|B \times C) = m_{X_K|X_M}(A|C) \qquad (16)$$

for any $A \subseteq \mathbf{X}_K$.

Remark. This is not the only way of generalization of the irrelevance concept, e.g. we could allow for conditioning by general sets and not only by rectangles on the left side of (16), i.e. the equality

$$m_{X_K|X_L X_M}(A|B) = m_{X_K|X_M}(A|B^{\downarrow M}) \qquad (17)$$

is satisfied for any $A \subseteq \mathbf{X}_K$ and $B \subseteq \mathbf{X}_{L\cup M}$. This definition is evidently more general, but it seemingly has little sense, as the "interesting" sets from the viewpoint of (conditional) independence are rectangles, or, more generally, joins.

Let us note that the conditioning in equalities (15) and (16) stands for an abstract conditioning rule (any of those mentioned in the previous section or some other [6]). Nevertheless, the validity of (15) and (16) may depend on the choice of conditioning rule. To demonstrate it let us present two simple examples.

[3]Let us note that somewhat weaker definition of irrelevance one can found in [2], where equality is substituted by proportionality. This notion has been later generalized using conjunctive combination rule [3].

Example 2 Let X_1 and X_2 be two binary variables (with values in $\mathbf{X}_i = \{a_i, \bar{a}_i\}$) with joint basic assignment m defined as follows:

$$\begin{aligned} m(\{(a_1, a_2)\}) &= \frac{1}{2}, \\ m(\mathbf{X}_1 \times \mathbf{X}_2 \setminus \{(\bar{a}_1, \bar{a}_2)\}) &= \frac{1}{4}, \\ m(\mathbf{X}_1 \times \mathbf{X}_2) &= \frac{1}{4}. \end{aligned}$$

From these values one can obtain

$$m^{\downarrow 2}(\{a_2\}) = m^{\downarrow 2}(\mathbf{X}_2) = \frac{1}{2},$$

and therefore

$$Bel^{\downarrow 2}(\{a_2\}) = \frac{1}{2}, \qquad Bel^{\downarrow 2}(\{\bar{a}_2\}) = 0.$$
$$Pl^{\downarrow 2}(\{a_2\}) = 1, \qquad Pl^{\downarrow 2}(\{\bar{a}_2\}) = \frac{1}{2}.$$

Computing conditional basic assignments (according to Dempster's conditioning rule) one can easily see that

$$\begin{aligned} m_{X_1|X_2}(\{a_1\}|\{a_2\}) &= m_{X_1|X_2}(\{a_1\}|\{\bar{a}_2\}) \\ &= \frac{1}{2} = m^{\downarrow 1}(\{a_1\}), \\ m_{X_1|X_2}(\{\bar{a}_1\}|\{a_2\}) &= m_{X_1|X_2}(\{\bar{a}_1\}|\{\bar{a}_2\}) \\ &= 0 = m^{\downarrow 1}(\{\bar{a}_1\}), \\ m_{X_1|X_2}(\mathbf{X}_1|\{a_2\}) &= m_{X_1|X_2}(\mathbf{X}_1|\{\bar{a}_2\}) \\ &= \frac{1}{2} = m^{\downarrow 1}(\mathbf{X}_1), \end{aligned}$$

i.e. X_1 and X_2 are irrelevant (with respect to Dempster's conditioning rule). On the other hand, as e.g.

$$\begin{aligned} & m_{X_1||X_2}((\{a_1\}||(\{a_2\})) \\ &= \frac{m(\{(a_1, a_2)\}))}{Bel(\{a_2\})} = 1 \neq \frac{1}{2} = m^{\downarrow 1}(\{a_1\}), \end{aligned}$$

they are not irrelevant with respect to focusing. ◇

Example 3 Let X_1 and X_2 be two binary variables (with values in $\mathbf{X}_i = \{a_i, \bar{a}_i\}$) with joint basic assignment m defined as follows:

$$\begin{aligned} m(\{(a_1, a_2)\}) &= \frac{1}{4}, \\ m(\{a_1\} \times \mathbf{X}_2) &= \frac{1}{4}, \\ m(\mathbf{X}_1 \times \{a_2\}) &= \frac{1}{4}, \\ m(\mathbf{X}_1 \times \mathbf{X}_2 \setminus \{(\bar{a}_1, \bar{a}_2)\}) &= \frac{1}{4}. \end{aligned}$$

From these values one can obtain

$$m^{\downarrow 2}(\{a_2\}) = m^{\downarrow 2}(\mathbf{X}_2) = \frac{1}{2},$$

and therefore

$$Bel^{\downarrow 2}(\{a_2\}) = \frac{1}{2}, \qquad Bel^{\downarrow 2}(\{\bar{a}_2\}) = 0,$$
$$Pl^{\downarrow 2}(\{a_2\}) = 1, \qquad Pl^{\downarrow 2}(\{\bar{a}_2\}) = \frac{1}{2}.$$

Evidently, it is not possible to condition by $\{\bar{a}_2\}$ and we have to confine ourselves to conditioning by $\{a_2\}$:

$$\begin{aligned} m_{X_1||X_2}(\{a_1\}||\{a_2\}) &= \tfrac{1}{2} &&= m^{\downarrow 1}(\{a_1\}), \\ m_{X_1||X_2}(\{\bar{a}_1\}||\{a_2\}) &= 0 &&= m^{\downarrow 1}(\{\bar{a}_1\}), \\ m_{X_1||X_2}(\mathbf{X}_1||\{a_2\}) &= \tfrac{1}{2} &&= m^{\downarrow 1}(\mathbf{X}_1), \end{aligned}$$

i.e. X_1 and X_2 are irrelevant (under focusing). On the other hand, as e.g.

$$\begin{aligned} & m_{X_1|X_2}((\{a_1\}|(\{\bar{a}_2\})) \\ &\quad = \frac{m(\{a_1\} \times \mathbf{X}_2) + m(\mathbf{X}_1 \times \mathbf{X}_2 \setminus \{(\bar{a}_1, \bar{a}_2)\})}{Pl(\{a_2\})} \\ &\quad = 1 \neq \frac{1}{2} = m^{\downarrow 1}(\{a_1\}), \end{aligned}$$

they are not irrelevant with respect to Dempster's conditioning rule. ♢

5 Relationship Between Independence and Irrelevance

As we demonstrated at the end of preceding section, different conditioning rules lead to different irrelevance concepts. Therefore we will study the relationships between independence and irrelevance separately for Dempster's conditioning rule and for focusing.

5.1 Dempster's rule of conditioning

For (unconditional) independence and irrelevance the following assertion holds true.

Theorem 2 *Let X_K and X_L ($K \cup L = \emptyset$) be independent groups of variables (under joint basic assignment m defined on $\mathbf{X}_{K\cup L}$). Then X_L are irrelevant to X_K with respect to Dempster's conditioning rule.*

Proof. Let X_K and X_L be independent. Then

$$m(A) = m^{\downarrow K}(A^{\downarrow K}) \cdot m^{\downarrow L}(A^{\downarrow L})$$

for any $A \subseteq \mathbf{X}_{K\cap L}$ for which $A = A^{\downarrow K} \times A^{\downarrow L}$, and $m(A) = 0$ otherwise, i.e. the only focal elements of m are rectangles. Therefore we have for arbitrary $A \subseteq \mathbf{X}_{K\cup L}$

$$\begin{aligned} Pl(A) &= \sum_{C: C\cap A \neq \emptyset} m(C) \\ &= \sum_{C: C\cap A \neq \emptyset} m^{\downarrow K}(C^{\downarrow K}) \cdot m^{\downarrow L}(C^{\downarrow L}) \\ &= \sum_{D: D\cap A^{\downarrow K} \neq \emptyset} m^{\downarrow K}(D) \cdot \sum_{E: E\cap A^{\downarrow L} \neq \emptyset} m^{\downarrow L}(E) \\ &= Pl^{\downarrow K}(A^{\downarrow K}) \cdot Pl^{\downarrow L}(A^{\downarrow L}). \end{aligned}$$

From this equality we immediately obtain that for all A such that $Pl^{\downarrow L}(A^{\downarrow L}) > 0$ equality

$$\frac{Pl(A)}{Pl^{\downarrow L}(A^{\downarrow L})} = Pl^{\downarrow K}(A^{\downarrow K})$$

is satisfied. But the left side of this equality is equal to $Pl_{X_K|X_L}(A^{\downarrow K}|A^{\downarrow L})$. As both conditional and marginal basic assignments can be obtained from corresponding plausibilities using the equality (4) and Möbius inversion (5), we immediately obtain that also for any fixed $B \subseteq \mathbf{X}_L$ such that $Pl^{\downarrow L}(B) > 0$

$$m_{K|L}(A|B) = m^{\downarrow K}(A)$$

for any $A \subseteq \mathbf{X}_K$, i.e. X_K and X_L are irrelevant. □

The reverse implication does not hold in general. To demonstrate it let us recall Example 2.

Example 2 *(Continued)* We have already shown that X_1 and X_2 are irrelevant (with respect to Dempster's conditioning rule). But they are not independent, as the focal elements are not rectangles, which contradicts Definition 2. ♢

Unfortunately, a generalization of Theorem 2 to conditional independence and conditional irrelevance does not hold, as can be seen from the following simple example.

Example 4 Let X_1, X_2 and X_3 be three variables with values in $\mathbf{X}_1, \mathbf{X}_2$ and $\mathbf{X}_3$ respectively, $\mathbf{X}_i = \{a_i, \bar{a}_i\}, i = 1, 2, 3$, and their joint basic assignment is defined as follows:

$$\begin{aligned} m(\{(x_1, x_2, x_3)\}) &= \tfrac{1}{16}, \\ m(\mathbf{X}_1 \times \mathbf{X}_2 \times \mathbf{X}_3) &= \tfrac{1}{2}, \end{aligned}$$

for $x_i = a_i, \bar{a}_i$, values of m on the remaining sets being 0, i.e. we have 9 focal elements — 8 singletons and the whole frame of discernment. Its marginal basic assignments on $\mathbf{X}_1 \times \mathbf{X}_3, \mathbf{X}_2 \times \mathbf{X}_3$ and $\mathbf{X}_3$ are

$$\begin{aligned} m^{\downarrow 13}(\{(x_1, x_3)\}) &= \tfrac{1}{8}, \\ m^{\downarrow 13}(\mathbf{X}_1 \times \mathbf{X}_3) &= \tfrac{1}{2}, \end{aligned}$$

$$\begin{aligned} m^{\downarrow 23}(\{(x_2, x_3)\}) &= \tfrac{1}{8}, \\ m^{\downarrow 23}(\mathbf{X}_2 \times \mathbf{X}_3) &= \tfrac{1}{2}, \end{aligned}$$

and

$$\begin{aligned} m^{\downarrow 3}(\{x_3\}) &= \tfrac{1}{4}, \\ m^{\downarrow 3}(\mathbf{X}_3) &= \tfrac{1}{2}, \end{aligned}$$

respectively (values of m of remaining subsets being 0, again). It is easy (but somewhat time-consuming) to show that for any $A \subseteq \mathbf{X}_1 \times \mathbf{X}_2 \times \mathbf{X}_3$ such that $A = A^{\downarrow 13} \bowtie A^{\downarrow 23}$

$$m(A) \cdot m^{\downarrow 3}(A^{\downarrow 3}) = m^{\downarrow 13}(A^{\downarrow 13}) \cdot m^{\downarrow 23}(A^{\downarrow 23}),$$

the values of remaining sets being zero, i.e. $\{1\} \perp\!\!\!\perp \{2\}|\{3\}\ [m]$ holds.

Now, let us show, that X_2 is not irrelevant to X_1 given X_3. To do so, we have to compute $m_{X_1|X_2X_3}$ and $m_{X_1|X_3}$. First, let us take into account that

$$Pl(\{(x_2, x_3)\}) = \frac{5}{8}$$

for any $x_i = a_i, \bar{a}_i, i = 2, 3$ and

$$Pl(\{x_3\}) = \frac{3}{4}$$

for both $x_3 = a_3, \bar{a}_3$ and that

$$(\mathbf{X}_1 \times \mathbf{X}_2 \times \mathbf{X}_3 \cap \{(a_2, a_3)\}^{\uparrow 123})^{\downarrow 1} = \mathbf{X}_1$$

and similarly

$$(\mathbf{X}_1 \times \mathbf{X}_3 \cap \{a_3\}^{\uparrow 13})^{\downarrow 1} = \mathbf{X}_1.$$

Then we have

$$m_{X_1|X_2X_3}(\{a_1\}|\{(a_2, a_3)\}) = \frac{m(\{(a_1, a_2, a_3)\})}{Pl(\{(a_2, a_3)\})} = \frac{1}{10},$$

$$m_{X_1|X_2X_3}(\{\bar{a}_1\}|\{(a_2, a_3)\}) = \frac{m(\{(\bar{a}_1, a_2, a_3)\})}{Pl(\{(a_2, a_3)\})} = \frac{1}{10},$$

$$m_{X_1|X_2X_3}(\mathbf{X}_1|\{(a_2, a_3)\}) = \frac{m(\mathbf{X}_1 \times \mathbf{X}_2 \times \mathbf{X}_3)}{Pl(\{(a_2, a_3)\})} = \frac{4}{5},$$

while

$$m_{X_1|X_3}(\{a_1\}|\{a_3\}) = \frac{m(\{(a_1, a_3)\})}{Pl(\{a_3\})} = \frac{1}{6},$$

$$m_{X_1|X_3}(\{\bar{a}_1\}|\{a_3\}) = \frac{m(\{(\bar{a}_1, a_3)\})}{Pl(\{a_3\})} = \frac{1}{6},$$

$$m_{X_1|X_3}(\mathbf{X}_1|\{a_3\}) = \frac{m(\mathbf{X}_1 \times \mathbf{X}_3)}{Pl(\{a_3\})} = \frac{2}{3},$$

i.e. $m_{X_1|X_2X_3} \neq m_{X_1|X_3}$. ♢

5.2 Focusing

In this subsection we will investigate mutual relationship between (conditional) independence and irrelevance based on the latter conditioning rule introduced in Section 3.

Theorem 3 *Let X_K and X_L ($K \cap L = \emptyset$) be independent groups of variables (under joint basic assignment m on $\mathbf{X}_{K \cup L}$). Then X_K and X_L are irrelevant with respect to focusing.*

Proof. Let X_K and X_L be independent. Then

$$m(A) = m^{\downarrow K}(A^{\downarrow K}) \cdot m^{\downarrow L}(A^{\downarrow L})$$

for any $A \subseteq \mathbf{X}_{K \cup L}$ for which $A = A^{\downarrow K} \times A^{\downarrow L}$, and $m(A) = 0$ otherwise, i.e. the only focal elements of m are rectangles. Therefore we have for arbitrary $A \subseteq \mathbf{X}_K$

$$\begin{aligned} Bel(A) &= \sum_{C \subseteq A} m(C) \\ &= \sum_{C \subseteq A} m^{\downarrow K}(C^{\downarrow K}) \cdot m^{\downarrow L}(C^{\downarrow L}) \\ &= \sum_{D: D \subseteq A^{\downarrow K}} m^{\downarrow K}(D) \cdot \sum_{E: E \subseteq A^{\downarrow L}} m^{\downarrow L}(E) \\ &= Bel^{\downarrow K}(A^{\downarrow K}) \cdot Bel^{\downarrow L}(A^{\downarrow L}). \end{aligned}$$

From this equality we immediately obtain that for all A such that $Bel^{\downarrow L}(A^{\downarrow L}) > 0$ equality

$$\frac{Bel(A)}{Bel^{\downarrow L}(A^{\downarrow L})} = Bel^{\downarrow K}(A^{\downarrow K})$$

is satisfied. But the left side of this equality is equal to $Bel_{X_K||X_L}(A^{\downarrow K}||A^{\downarrow L})$. As the both conditional and marginal basic assignments can be obtained from corresponding beliefs using Möbius inversion (5) we immediately obtain that also for any fixed $B \subseteq \mathbf{X}_L$ such that $Bel(B) > 0$

$$m_{X_K||X_L}(A||B) = m^{\downarrow K}(A)$$

for any $A \subseteq \mathbf{X}_K$, i.e. X_K and X_L are irrelevant. □

The reverse implication does not hold again, as can be seen from the following simple example (continuation of Example 3).

Example 3 *(Continued)* We have already proven that X_1 and X_2 are irrelevant (under focusing). But they are not independent, as the focal elements are not rectangles, which again contradicts Definition 2.♢

Up to now the results presented in this subsection have been exactly the same as in the preceding one.

Now, let us study the problem of the relationship between conditional independence and irrelevance. For this purpose, let us recall Example 4.

Example 4 *(Continued)* We have already shown that although X_1 and X_2 are conditionally independent given X_3, X_2 is not irrelevant to X_1 given X_3 under Dempster's rule of conditioning.

Now, let us us check whether X_2 is irrelevant to X_1 given X_3 under focusing. To do so, we have to compute $m_{X_1||X_2X_3}$ and $m_{X_1||X_3}$. Again, we have to take into account that

$$Bel(\{(x_2, x_3)\}) = \frac{1}{8}$$

for any $x_i = a_i, \bar{a}_i, i = 2, 3$ and

$$Bel(\{x_3\}) = \frac{1}{4}$$

for both $x_3 = a_3, \bar{a}_3$ and the fact that there does not exist any focal element A of m such that $A \subseteq \{(x_2, x_3)\}^{\uparrow 123}$ (for any pair (x_2, x_3)) and $A^{\downarrow 1} = \mathbf{X}_1$ and similarly there does not exist any focal element B of $m^{\downarrow 13}$ such that $B \subseteq \{x_3\}^{\uparrow 13}$ (for any x_3) and $B^{\downarrow 1} = \mathbf{X}_1$. Therefore we have

$$\begin{aligned} & m_{X_1||X_2X_3}(\{a_1\}||\{(x_2, x_3)\}) \\ & \qquad = \frac{m(\{(a_1, x_2, x_3)\})}{Bel(\{(x_2, x_3)\})} = \frac{1}{2}, \\ & m_{X_1||X_2X_3}(\{\bar{a}_1\}||\{(x_2, x_3)\}) \\ & \qquad = \frac{m(\{(\bar{a}_1, x_2, x_3)\})}{Bel(\{(x_2, x_3)\})} = \frac{1}{2}, \\ & m_{X_1||X_2X_3}(\mathbf{X}_1||\{(x_2, x_3)\}) = 0, \end{aligned}$$

for any pair $(x_2, x_3) \in \mathbf{X}_2 \times \mathbf{X}_3$ and

$$\begin{aligned} m_{X_1||X_3}(\{a_1\}||\{x_3\}) &= \frac{m(\{(a_1, x_3)\})}{Bel(\{x_3\})} = \frac{1}{2}, \\ m_{X_1||X_3}(\{\bar{a}_1\}||\{x_3\}) &= \frac{m(\{(\bar{a}_1, x_3)\})}{Bel(\{x_3\})} = \frac{1}{2}, \\ m_{X_1||X_3}(\mathbf{X}_1||x_3) &= 0, \end{aligned}$$

for any $x_3 \in \mathbf{X}_3$, i.e. $m_{X_1||X_2X_3} = m_{X_1||X_3}$ when conditioning by singletons, which is quite different from the previous case, based on Dempster's conditioning rule.

Nevertheless, to demonstrate that X_2 is irrelevant to X_1 given X_3 we have also to check the validity of equality (16) for a general rectangle $B \times C$ such that $Bel(B \times C) > 0$. As both X_2 and X_3 are binary, only three situations may happen:

$B = \mathbf{X}_2$ and $C = \mathbf{X}_3$: in this case equality (16) is trivially satisfied, as conditional basic assignments on both sides are, in fact, marginal basic assignments on X_1, and therefore identical;

$B = \mathbf{X}_2$ and $C = \{x_3\}$ for $x_3 = a_3, \bar{a}_3$: in this case equality (16) is again satisfied, as conditional basic assignment on the left side is, in fact, the same as that on the right side;

$B = \{x_2\}$ for $x_2 = a_2, \bar{a}_2$ and $C = \mathbf{X}_3$: this is the nontrivial case, corresponding to unconditional irrelevance (15); nevertheless, its validity need not be checked, since X_1 and X_2 are not (unconditionally) independent, as can be easily checked.

Therefore X_2 is irrelevant to X_1 given X_3 (under focusing). ♢

Let us finish the section with a partial generalization of Theorem 3, which, maybe surprisingly, proves that conditioning by sets which are not rectangles is sensible.

Theorem 4 *Let X_K and X_L be conditionally independent groups of variables given X_M under joint basic assignment m on $\mathbf{X}_{K \cup L \cup M}$ (K, L, M disjoint, $K \neq \emptyset \neq L$). Then*

$$m_{X_K||X_LX_M}(A||B) = m_{X_K||X_M}(A||B^{\downarrow M}) \qquad (18)$$

for any $m^{\downarrow L \cup M}$-atom $B \subseteq \mathbf{X}_{L \cup M}$ such that $B^{\downarrow M}$ is $m^{\downarrow M}$-atom and $A \subseteq \mathbf{X}_K$.

Proof. Let X_K and X_L be conditionally independent given X_M. Then

$$\begin{aligned} & m(C) \cdot m^{\downarrow M}(C^{\downarrow M}) \\ & \qquad = m^{\downarrow K \cup M}(C^{\downarrow K \cup M}) \cdot m^{\downarrow L \cup M}(C^{\downarrow L \cup M}) \end{aligned}$$

holds for any $C \subseteq \mathbf{X}_{K \cup L \cup M}$ such that $C = C^{\downarrow K \cup M} \bowtie C^{\downarrow L \cup M}$, and $m(C) = 0$ otherwise. From this equality we immediately obtain that for all C such that $m^{\downarrow L}(C^{\downarrow L}) > 0$ equality

$$\frac{m(C)}{m^{\downarrow L \cup M}(C^{\downarrow L \cup M})} = \frac{m^{\downarrow K \cup M}(C^{\downarrow K \cup M})}{m^{\downarrow M}(C^{\downarrow M})}$$

is satisfied. If $C^{\downarrow L \cup M}$ is an atom, then $m^{\downarrow L \cup M}(C^{\downarrow L \cup M}) = Bel^{\downarrow L \cup M}(C^{\downarrow L \cup M})$ (and analogously $m^{\downarrow M}(C^{\downarrow M}) = Bel^{\downarrow M}(C^{\downarrow M})$ if $C^{\downarrow M}$ is an atom) and this equality may be rewritten into the form

$$\frac{m(C)}{Bel^{\downarrow L \cup M}(C^{\downarrow L \cup M})} = \frac{m^{\downarrow K \cup M}(C^{\downarrow K \cup M})}{Bel^{\downarrow M}(C^{\downarrow M})}.$$

If we denote $C^{\downarrow L \cup M}$ by B, we obtain

$$m_{X_K||X_LX_M}(C^{\downarrow K}||B) = m_{X_K||X_M}(C^{\downarrow K}||B^{\downarrow M}).$$

If $A \neq C^{\downarrow K}$, then $m(A^{\uparrow K \cup L \cup M} \cap B^{\uparrow K \cup L \cup M}) = 0$ and therefore equality (18) is trivially satisfied. □

From this theorem it is evident, that conditions under which conditional independence implies conditional irrelevance are rather restrictive.

The requirement in Theorem 4 for B being an atom is substantial, as can be seen from the following simple example (again continuation of Example 4).

Example 4 *(Continued)* Let us consider a set $B = \{(a_2, a_3), (\bar{a}_2, \bar{a}_3)\} \subseteq \mathbf{X}_2 \times \mathbf{X}_3$. One can easily compute that $Bel(B) = \frac{1}{4}$ and therefore

$$\begin{aligned} & m_{X_1|X_2X_3}(\{a_1\}|B) \\ & \qquad = \frac{m(\{a_1, a_2, a_3\}) + m(\{a_1, \bar{a}_2, \bar{a}_3\})}{Bel(B)} = \frac{1}{2}, \end{aligned}$$

$$m_{X_1|X_2X_3}(\{\bar{a}_1\}|B) = \frac{m(\{\bar{a}_1, a_2, a_3\}) + m(\{\bar{a}_1, \bar{a}_2, \bar{a}_3\})}{Bel(B)} = \frac{1}{2},$$
$$m_{X_1|X_2X_3}(\mathbf{X}_1|B) = 0,$$

while,

$$m_{X_1|X_3}(\{a_1\}|B^{\downarrow 3}) = m^{\downarrow 1}(\{a_1\}) = \frac{1}{4},$$
$$m_{X_1|X_3}(\{\bar{a}_1\}|B^{\downarrow 3}) = m^{\downarrow 1}(\{\bar{a}_1\}) = \frac{1}{4},$$
$$m_{X_1|X_2X_3}(\mathbf{X}_1|B^{\downarrow 3}) = m^{\downarrow 1}\ (\mathbf{X}_1) = \frac{1}{2}.$$

as $B^{\downarrow 3} = \mathbf{X}_3$. ♢

6 Conclusions

We presented two conditional rules for basic assignment and studied the relationship between (conditional) independence and (conditional) irrelevance (based on these conditioning rules) in evidence theory.

While in unconditional case independence implies irrelevance and not vice versa (as expected), for conditional independence such an implication does not hold, in general. Therefore, it is necessary to be cautious when constructing Bayesian-network-like models in evidence theory, as the mutual relationship is more complicated than in probabilistic framework.

It may be of some interest to study another way of conditioning presented in [8], however, its application will be more complicated, as conditional basic assignment must be obtained via Möbius transform from conditional beliefs. Furthermore, we are somewhat sceptic about the result.

Acknowledgements

The work of the author was supported by the grant GA ČR 201/09/1891. The author would like to express her gratitude to both referees for their inspiring comments.

References

[1] C. Beeri, R. Fagin, D. Maier, M. Yannakakis, On the desirability of acyclic database schemes, *J. of the Association for Computing Machinery,* **30** (1983), 479–513.

[2] B. Ben Yaghlane, Ph. Smets and K. Mellouli, Belief functions independence: I. the marginal case. *Int. J. Approx. Reasoning*, **29** (2002), 47–70.

[3] B. Ben Yaghlane, Ph. Smets and K. Mellouli, Belief functions independence: II. the conditional case. *Int. J. Approx. Reasoning*, **31** (2002), 31-75.

[4] B. Ben Yaghlane, Ph. Smets and K. Mellouli, Directed evidential networks with conditional belief functions. *Proceedings of ECSQARU 2003,* eds. T. D. Nielsen, N. L. Zhang, 291–305.

[5] I. Couso, S. Moral and P. Walley, Examples of independence for imprecise probabilities, *Proceedings of ISIPTA'99,* eds. G. de Cooman, F. G. Cozman, S. Moral, P. Walley, 121–130.

[6] M. Daniel, Belief conditioning Rules for classic belief functions, *Proceedings of WUPES'09,* eds. T. Kroupa, J. Vejnarová, J., 46–56.

[7] G. de Cooman, Possibility theory I – III. *Int. J. General Systems* **25** (1997), pp. 291–371.

[8] R. Fagin and J. Y. Halpern, A new approach to updating beliefs, *Uncertainty in artifiacial intelligence,* eds. Bonissone et al., vol. VI, pp. 347–374, Elsevier, 1991.

[9] R. Jiroušek and J. Vejnarová, Compositional models and conditional independence in Evidence Theory, *Int. J. Approx. Reasoning,* **52** (2011), 316-334.

[10] S. L. Lauritzen, *Graphical Models.* Oxford University Press, 1996.

[11] G. Shafer, *A Mathematical Theory of Evidence.* Princeton University Press, Princeton, New Jersey, 1976.

[12] P. P. Shenoy, Conditional independence in valuation-based systems. *Int. J. Approx. Reasoning,* **10** (1994), 203-234.

[13] J. Vejnarová, Conditional independence relations in possibility theory. *Int. J. Uncertainty, Fuzziness and Knowledge-Based Systems,* vol. 8, pp. 253–269, 2000.

[14] J. Vejnarová, On two notions of independence in evidence theory, *Proceedings of 11th Czech-Japan Seminar on Data Analysis and Decision Making under Uncertainty*, eds. T. Itoh, A. Shirouha, Sendai University, 2008, pp. 69–74.

[15] J. Vejnarová, On conditional independence in evidence theory, *Proceedings of ISIPTA'09,* eds. T. Augustin, F. P. A. Coolen, S. Moral and M. C. M. Troffaes,Durham, UK, 2009, pp. 431–440.

[16] J. Vejnarová, A thorough comparison of two conditional independence concepts for belief functions, *Proceedings of Workshop on the Theory of Belief Functions,* Brest, France, 2010.