Conditional Independence in Evidence Theory

Jiřina Vejnarová
Institute of Information Theory and Automation
Academy of Sciences of the Czech Republic
&
University of Economics, Prague
vejnar@utia.cas.cz

Abstract

We overview the existing conditional independence concepts from the viewpoint of their relation to possibility theory. We will show that a suitable notion — from this perspective — seems to be strong independence in the wider framework of credal sets.

Keywords: Conditional independence, Evidence theory, Probability theory, Possibility theory, Credal sets.

1 Introduction

The complexity of practical problems that are of primary interest in the field of artificial intelligence usually results in the necessity to construct models with the aid of a great number of variables: more precisely, hundreds or thousands rather than tens. Processing distributions of such dimensionality would not be possible without some tools allowing us to reduce demands on computer memory. Conditional independence, which belongs among such tools, allows the expression of these multidimensional distributions by means of low-dimensional ones, and therefore to substantially decrease demands on computer memory.

For three centuries, probability theory has been the only mathematical tool at our disposal for uncertainty quantification and processing. As a result, many important theoretical and practical advances have been achieved in this field. However, during the last forty years some new mathematical tools have emerged as alternatives to probability theory. They are used in situations whose nature of uncertainty does not meet the requirements of probability theory, or those in which probabilistic approaches employ criteria that are too strict. Nevertheless, probability theory has always served as a source of inspiration for the development of these non-probabilistic calculi and these calculi have been continually confronted with probability theory and mathematical statistics from various points of view. Good examples of this fact include the numerous papers studying conditional independence in various calculi [2, 3, 4, 5, 6, 7, 9, 11, 12, 13, 14, 16].

With this contribution, we will concentrate ourselves to evidence theory, which can be viewed as a generalization of both probability and possibility theories. After a brief review of basic notation and terminology necessary for understanding the next part of the paper (Section 2) we will summarize in Section 3 the notions of conditional independence in evidence theory proposed by various authors, and show that although all these notions are generalizations the notion of conditional independence in probability theory, nothing similar holds for possibility theory. In Section 4 we will suggest how this problem can be solved within a wider framework of credal sets and in Section 5 we will study the relation of strong independence to conditional T-independence thoroughly studied in [16].

L. Magdalena, M. Ojeda-Aciego, J.L. Verdegay (eds): Proceedings of IPMU'08, pp. 938–945
Torremolinos (Málaga), June 22–27, 2008

2 Basic Notions

Consider a finite index set $N = \{1, 2, \ldots, n\}$ and finite sets $\{\mathbf{X}_i\}_{i \in N}$. In this text we will consider *multidimensional frame of discernment*

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

and its *subframes*. For $K \subset N$, $\mathbf{X}_K$ denotes a Cartesian product of those $\mathbf{X}_i$, for which $i \in K$:

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

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 $K = \{i_1, i_2, \ldots, i_\ell\}$

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

Analogously, for $K \subset L \subseteq N$ and $A \subset \mathbf{X}_L$, $A^{\downarrow K}$ will denote a *projection* of A into $\mathbf{X}_K$:

$$A^{\downarrow K} = \{y \in \mathbf{X}_K | \exists x \in A : y = x^{\downarrow K}\}.$$

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

In addition to the projection, in this text we will need also the opposite operation which will be called extension. By an *extension* of two sets $A \subseteq \mathbf{X}_K$ and $B \subseteq \mathbf{X}_L$ we will understand a set

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

Let us note that if K and L are disjoint, then

$$A \otimes B = A \times B.$$

Consider a *basic (probability or belief) assignment* (or just assignment) m on $\mathbf{X}_N$, i.e.

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

for which $\sum_{A \subseteq \mathbf{X}_N} m(A) = 1$. For each $K \subset N$ its *marginal basic assignment* is defined (for each $B \subseteq \mathbf{X}_K$):

$$m^{\downarrow K}(B) = \sum_{A \subseteq \mathbf{X}_N : A^{\downarrow K} = B} m(A).$$

Having two basic assignments m_1 and m_2 on $\mathbf{X}_K$ and $\mathbf{X}_L$, respectively (we assume that $K, L \subseteq N$), we say that these assignments are *projective* if

$$m_1^{\downarrow K \cap L} = m_2^{\downarrow K \cap L},$$

which occurs if and only if there exists a basic assignment m on $\mathbf{X}_{K \cup L}$ such that both m_1 and m_2 are marginal assignments of m.

Given a basic assignment m we can obtain *belief, plausibility* and *commonality functions* via the following formulae:

$$\begin{aligned} Bel(A) &= \sum_{B \subseteq A} m(B); \\ Pl(A) &= \sum_{B \cap A \neq \emptyset} m(B); \\ Q(A) &= \sum_{A \subseteq B \subseteq \mathbf{X}_N} m(B). \end{aligned}$$

Now let us concentrate our attention to two special cases of basic assignments.

A basic assignment is called *Bayesian* if all its focal elements[1] are singletons. In this case $Bel(A) = Pl(A) = P(A)$, called a *probability measure*, and $Q(A) = m(A)$ for all $A \in \mathcal{P}(X_N)$; m can be substituted by a point function

$$p : \mathbf{X}_N \longrightarrow [0, 1]$$

called a *probability distribution*.

A body of evidence[2] is called *consonant* if its focal elements are nested. In this case

$$Pl(A) = \max_{B \subseteq A} Pl(B),$$

i.e. plausibility function becomes a *possibility measure* Π and its values for any $A \subseteq \mathbf{X}_N$ can be obtained from a *possibility distribution*

$$\pi : \mathbf{X}_N \longrightarrow [0, 1]$$

via the formula:

$$\Pi(A) = \max_{x \in A} \pi(x).$$

Connection between basic assignment m and a possibility distribution π is expressed by the following formula:

$$\pi(x) = \sum_{x \in A \in \mathcal{P}(\mathbf{X}_N)} m(A).$$

[1] A set $A \in \mathcal{P}(X_N)$ is a *focal element* if $m(A) > 0$.

[2] A *body of evidence* is a pair $(\mathcal{F}, m)$, where $\mathcal{F}$ is the set of all focal elements.

3 Conditional Independence Concepts

In this section we will overview different conditional independence notions introduced in the framework of evidence theory. Before doing that let us stress, that we are interested in the conditional independence from the viewpoint of *decomposition* of multidimensional models and not in its behavioral interpretation.

3.1 Conditional Non-Interactivity

This notion was introduced by Ben Yaghlane *et al.* [2], but their definition is, in fact, equivalent to the conditional independence notion by Shenoy and Studený [12, 14].

Let X, Y and Z be variables taking their values in $\mathbf{X}, \mathbf{Y}$ and $\mathbf{Z}$, respectively, and m be a joint basic assignment on $\mathcal{P}(\mathbf{X} \times \mathbf{Y} \times \mathbf{Z})$. Variables X and Y are *conditionally non-interactive*[3] given Z with respect to m iff the equality

$$Q(A) \cdot Q^{\downarrow Z}(A^{\downarrow Z}) = Q^{\downarrow XZ}(A^{\downarrow XZ}) \cdot Q^{\downarrow YZ}(A^{\downarrow YZ}) \quad (1)$$

holds for any $A \in \mathcal{P}(\mathbf{X} \times \mathbf{Y} \times \mathbf{Z})$.

The authors proved that conditional non-interactivity satisfies so-called semigraphoid properties, usually taken as sound properties of a conditional independence relation.

The authors also claim that their definition of conditional non-interactivity is equivalent to that of strong conditional independence introduced by Almond [1], which is based on Dempster's rule.

Nevertheless, this notion of independence does not seem to be appropriate, as it is *not consistent with marginalization.* What does it mean can be seen from the following simple example (by Studený).

Example 1 Let X, Y and Z be three binary variables $\mathbf{X} = \mathbf{Y} = \mathbf{Z} = \{u, v\}$ and m_{XZ} and m_{YZ} two basic assignments, both of them having only two focal elements:

$$\begin{aligned} m_{XZ}(\{(u,v),(v,v)\}) &= .5,\\ m_{XZ}(\{(u,v),(v,u)\}) &= .5,\\ m_{YZ}(\{(u,v),(v,v)\}) &= .5,\\ m_{YZ}(\{(u,v),(v,u)\}) &= .5. \end{aligned}$$

Since their marginals are projective

$$\begin{aligned} m_{XZ}^{\downarrow Z}(\{v\}) = m_{YZ}^{\downarrow Z}(\{v\}) &= .5,\\ m_{XZ}^{\downarrow Z}(\{u,v\}) = m_{YZ}^{\downarrow Z}(\{u,v\}) &= .5, \end{aligned}$$

there exist an extension of both of them. Nevertheless, the application of the equality (1) to these basic assignments leads to the following values of the joint "basic assignment":

$$\begin{aligned} m_{NI}(\mathbf{X} \times \mathbf{Y} \times \{v\}) &= .25,\\ m_{NI}(\mathbf{X} \times \{u\} \times \{v\}) &= .25,\\ m_{NI}(\{u\} \times \mathbf{Y} \times \{v\}) &= .25,\\ m_{NI}(\{(u,u,v),(v,v,u)\}) &= .5,\\ m_{NI}(\{(u,u,v)\}) &= -.25, \end{aligned}$$

i.e. we ore out of evidence theory. $\diamondsuit$

From the viewpoint of multidimensional models it seems to be a substantial drawback.

3.2 Factorization

Another definition of conditional independence was recently proposed by Jiroušek [9]. His definition is based on the notion of the *operator of composition* introduced in the evidence theory in [10] as follows:

For two arbitrary basic assignments m_1 on $\mathbf{X}_K$ and m_2 on $\mathbf{X}_L$ a *composition* $m_1 \triangleright m_2$ is defined for all $C \subseteq \mathbf{X}_{K \cup L}$ by one of the following expressions:

[a] if $m_2^{\downarrow K \cap L}(C^{\downarrow K \cap L}) > 0$ and $C = C^{\downarrow K} \otimes C^{\downarrow L}$ then

$$(m_1 \triangleright m_2)(C) = \frac{m_1(C^{\downarrow K}) \cdot m_2(C^{\downarrow L})}{m_2^{\downarrow K \cap L}(C^{\downarrow K \cap L})};$$

[b] if $m_2^{\downarrow K \cap L}(C^{\downarrow K \cap L}) = 0$ and $C = C^{\downarrow K} \times \mathbf{X}_{L \setminus K}$ then

$$(m_1 \triangleright m_2)(C) = m_1(C^{\downarrow K});$$

[3]Let us note that the definition presented in [2] is based on conjunctive Dempster's rule, but the authors proved its equivalence with 1.

[**c**] in all other cases

$$(m_1 \triangleright m_2)(C) = 0.$$

The relation of above defined operator of composition to its probabilistic pre-image is expressed by the following simple lemma proven in [10].

Lemma 1 *Let m_1 and m_2 be Bayesian basic assignments on $\mathbf{X}_K$ and $\mathbf{X}_L$, respectively, for which*

$$m_2^{\downarrow K\cap L}(A) = 0 \implies m_1^{\downarrow K\cap L}(A) = 0 \quad (2)$$

for any $A \subseteq \mathbf{X}_{K\cup L}$. Then $m_1 \triangleright m_2$ is a Bayesian basic assignment.

The ternary relation of *factorization* is defined (for disjoined $I, J, K \subset N, I \neq \emptyset \neq J$) in the following way

$$X_I \perp\!\!\!\perp_m X_J | X_K \iff m^{\downarrow I\cup J\cup K} = m^{\downarrow I\cup K} \triangleright m^{\downarrow J\cup K}.$$

It was shown in [9] that this relation also satisfies semigraphoid properties, but, in contrary to conditional non-interactivity, it does not suffer from the drawback of inconsistency with marginalization, as can be seen from the following example.

Example 1 *(Continued)* Taking into consideration the results obtained in the previous part, one can easily realize that using the above presented definition of the operator of composition we will obtain the following joint assignment (under factorization):

$$m_F(\mathbf{X} \times \mathbf{Y} \times \{v\}) = .5$$
$$m_F(\{(u,u,v),(v,v,u)\}) = .5 \qquad \diamondsuit$$

3.3 Relation to Probability and Possibility Theories

Evidence theory can be understood as a generalization of both probability and possibility theories. From this perspective, one could expect, that the (conditional) independence notion should be a generalization of (conditional) independence notions of both theories.

X_1	0	1
π_1	1	.7

X_2	0	1
π_2	.5	1

Table 1: Possibility distributions π_1 and π_2.

It is valid for both concepts (non-interactivity and factorization) for probability theory. As concerns conditional non-interactivity, it can be seen from the fact, that commonality function of Bayesian basic assignment equals this basic assignment. In that case (1) becomes

$$p(x,y,z) \cdot p(z) = p(x,z) \cdot p(y,z)$$

for all $(x,y,z) \in \mathbf{X} \times \mathbf{Y} \times \mathbf{Z}$.

For factorization concept it follows from Lemma 1 and the following lemma taken from [8].

Lemma 2 *Let I, J, K be disjoint subsets of N, I and J be nonempty. For a probability distribution p defined on $\mathbf{X}_N$*

$$p^{\downarrow I\cup J\cup K} = p^{\downarrow I\cup K} \triangleright p^{\downarrow J\cup K}$$

if and only if X_I and X_J are conditionally independent given X_K with respect to p.

Unfortunately, it does not hold for possibility theory. It is even worse, having two independent consonant bodies of evidence, the resulting body of evidence is not consonant. Let us demonstrate it by the following simple example. Before doing that let us note that both conditional independence concepts presented above collapse, if $K = \emptyset$, to

$$m(A) = m^{\downarrow I}(B) \cdot m^{\downarrow J}(C) \qquad (3)$$

for all $A = B \times C$ and $m(A) = 0$ otherwise.

Example 2 Let $\mathbf{X}_1 = \mathbf{X}_2 = \{0,1\}$ be two frames of discernment and π_1, π_2 defined by Table 1 be possibility distributions defined on them.

From these marginal possibilities we can get basic assignments contained in Table 2 and (under the independence assumption) via formula (3) the joint assignment in Table 3. It is evident that the focal elements are not nested. $\diamondsuit$

A_1	$\{0\}$	$\mathbf{X}_1$
m_1	.3	.7

A_2	$\{1\}$	$\mathbf{X}_2$
m_2	.5	.5

Table 2: Corresponding basic assignments m_1 and m_2.

$A_1 \times A_2$	$m_1 \cdot m_2$
$\{0\} \times \{1\}$	.15
$\{0\} \times \mathbf{X}_2$	.15
$\mathbf{X}_1 \times \{1\}$	.35
$\mathbf{X}_1 \times \mathbf{X}_2$	.35

Table 3: Resulting basic assignments on $\mathbf{X}_1 \times \mathbf{X}_2$.

4 Strong Independence

This problem can be avoided if we take into account the fact that both evidence and possibility theories can be considered as special kinds of imprecise probabilities. Let us start with the definition of the fundamental concept of a credal set.

A *credal set* $\mathcal{M}(X)$ about a variable X is defined as a set of probability measures about the values of this variable. In order to simplify the expression of operations with credal sets, it is often considered [11] that a credal set is the set of probability distributions associated to the probability measures in it. Under such consideration a credal set can be expressed as a *convex hull* of its extreme distributions

$$\mathcal{M}(X) = \mathrm{CH}\{\mathrm{ext}(\mathcal{M}(X))\}.$$

Again, there exist numerous definitions of independence for credal sets [6], but we have chosen strong independence, as it seems to be most proper for multidimensional models.

We say that X and Y are *strongly independent* with respect to $\mathcal{M}(XY)$ iff (in terms of probability distributions)

$$\mathcal{M}(XY) = \mathrm{CH}\{p_1 \cdot p_2 : p_1 \in \mathcal{M}(X), p_2 \in \mathcal{M}(Y)\}. \quad (4)$$

Note that any possibility distribution can be associated with a credal set

$$\mathcal{M}(\pi) = \{p : p(x) \leq \pi(x), x \in \mathbf{X}\}.$$

π_S X_2	0	1
$X_1 = 0$	.5	1
$X_1 = 1$	.35	.7

Table 4: Joint possibility distribution obtained by strong independence

We can utilize this fact to find an extension under strong independence of the distributions in Example 2.

Example 2 *(Continued)* Possibility distribution π_1 can be associated with the credal set

$$\begin{aligned}\mathcal{M}(\pi_1) &= \{p_1 : p_1(1) \leq 0.7, p_1(0) = 1 - p_1(1)\}\\ &= \mathrm{CH}\{(0.3, 0.7), (1, 0)\}\end{aligned}$$

and similarly π_2 to the set

$$\begin{aligned}\mathcal{M}(\pi_2) &= \{p_2 : p_2(0) \leq 0.5, p_2(1) = 1 - p_2(1)\}\\ &= \mathrm{CH}\{(0.5, 0.5), (0, 1)\}.\end{aligned}$$

The set $\mathcal{M}(\pi_1 \cdot \pi_2)$ can be then obtained via formula (4):[4]

$$\begin{aligned}\mathcal{M}(\pi_1 \cdot \pi_2) = \mathrm{CH}\{&(.15, .15, .35, .35), (.5, .5, 0, 0),\\ &(0, .3, 0, .7), (0, 1, 0, 0)\}.\end{aligned}$$

As suprema must be achieved in extreme points, one can easily obtain the joint possibility distribution π_S in Table 4. ♢

There exist several generalizations of this notion to conditional independence, see e.g. [11], but the following definition is suggested by the authors as the most appropriate for the marginal problem, which is in the center of our attention for a long time.

Given three variables X, Y and Z we say that X and Y are *independent on the distribution* given Z under global set $\mathcal{M}(X, Y, Z)$ iff

$$\begin{aligned}\mathcal{M}(X, Y, Z) = \{(p_1 \cdot p_2)/p_1^{\downarrow Z} : p_1 \in \mathcal{M}(X, Z),\\ p_2 \in \mathcal{M}(X, Z), p_1^{\downarrow Z} = p_2^{\downarrow Z}\}.\end{aligned}$$

It was proven by Moral and Cano [11] that strong independence on distribution satisfies

[4] Let us note that the extreme points are $(p(0,0), p(0,1), p(1,0), p(1,1))$.

semigraphoid properties. Therefore, from this formal point of view it is comparable with the notions of conditional non-interactivity and factorization presented in Section 3.

It seems to be interesting how Example 1 can be solved by strong independence.

Example 1 *(Continued)* From the values of the basic assignments m_{XZ} and m_{YZ} we will obtain

$$\begin{array}{c} p_{XZ}(u,u) = 0 = p_{YZ}(u,u) \\ p_{XZ}(u,v) \in [0,1] \ni p_{YZ}(u,v), \\ p_{XZ}(v,u) \in [0,.5] \ni p_{YZ}(v,u), \\ p_{XZ}(v,v) \in [0,.5] \ni p_{XZ}(v,v). \end{array}$$

Credal set $\mathcal{M}(XZ)$ is therefore

$$\begin{aligned} \mathcal{M}(XZ) \;=\; & \mathrm{CH}\{(0,1,0,0),(0,.5,0,.5) \\ & (0,.5,0.5,0),(0,0,.5,.5)\} \end{aligned}$$

and credal set $\mathcal{M}(YZ)$ is identical. We can see, that the first two probability distributions are projective and the remaining two as well. Therefore under the assumption of strong conditional independence we will get the following joint credal set

$$\begin{aligned} \mathcal{M}(XYZ) = \mathrm{CH}\{ & (0,1,0,0,0,0,0,0), \\ & (0,.5,0,.5,0,0,0,0), \\ & (0,.5,0,0,0,.5,0,0), \\ & (0,.25,0,.25,0,.25,0,.25), \\ & (0,.5,0,0,0,0,.5,0), \\ & (0,.0,0,.5,0,0,.5,0), \\ & (0,0,0,0,0,.5,.5,0), \\ & (0,0,0,0,0,0,.5,.5)\}. \end{aligned}$$

From $\mathcal{M}(XYZ)$ we can easily get values of upper and lower probabilities of all singletons as well as values of bigger subsets. For example, for the focal elements obtained by the operator of composition $\triangleright$ we have

$$\begin{array}{l} P(\mathbf{X} \times \mathbf{Y} \times \{v\}) \in [.5,1], \\ P(\{(u,u,v),(v,v,u)\}) \in [.5,1], \end{array}$$

which perfectly coincides with the results obtained under factorization. It is easy but rather tiresome to check that it holds for all 255 nonempty subsets of $\{u,v\}^3$. $\diamondsuit$

5 Relation to T-independence

In this section we will recall the notion of possibilistic T-independence introduced in [15] and studied in more details in [16, 17] and study its relation to strong conditional independence in the framework of possibility theory.

Before doing that let us recall the notion of a triangular norm, because T-independence is parameterized by it.

A *triangular norm* (or a *t-norm*) T is an isotonic, associative and commutative binary operator on $[0,1]$ (i.e. $T : [0,1]^2 \to [0,1]$) satisfying the boundary condition: for any $x \in [0,1]$

$$T(1,x) = x.$$

A t-norm T is called *continuous* if T is a continuous function.

Here are the most important examples of continuous t-norms:

(i) *Gödel's t-norm:* $T_G(a,b) = \min(a,b)$;

(ii) *product t-norm:* $T_\Pi(a,b) = a \cdot b$;

(iii) *Łukasziewicz's t-norm:* $T_L(a,b) = \max(0, a+b-1)$.

Given a possibility measure Π on $\mathbf{X} \times \mathbf{Y} \times \mathbf{Z}$ with the respective distribution $\pi(x,y,z)$, variables X and Y are *possibilistically conditionally T-independent*[5] given Z if, for any pair $(x,y) \in \mathbf{X} \times \mathbf{Y}$,

$$\begin{aligned} & \pi_{XYZ}(x,y,z) \\ & \quad = T(\pi_{X|_T Z}(x|_T z), \pi_{YZ}(y,z)), \end{aligned} \qquad (5)$$

where $\pi_{X|_T Z}(x|_T z)$ is defined as *any* solution of the equation

$$\pi(x,z) = T(\pi(z), \pi(x|z)).$$

For more details see on conditional T-independence see [16].

[5] Let us note that the definition presented in [15] is a bit different, but (5) is one of its equivalent characterizations.

π_T X_2	0	1
$X_1 = 0$	.5	1
$X_1 = 1$	$T(0.5, 0.7)$	.7

Table 5: Joint possibility distribution obtained by T-independence

To get a basic idea about this notion (although in the unconditional case), let us apply it to the possibility distributions from Example 2.

Example 2 *(Continued)* Under the assumption of T-independence we will obtain using (5) the joint distribution contained in Table 5. One can easily see that the values of π_T are identical for different t-norms with the exception of $\pi_T(1, 0)$, which is the only one depending on its choice. For product t-norm, we will get the same result as for strong independence (cf. Table 4). ♢

One can hardly expect that conditional strong independence is equivalent to conditional product-independence. Nevertheless, the following assertion expresses their relationship.

Theorem 1 *Let X and Y are strongly conditionally independent in distribution given Z. Then X and Y are conditionally product-independent.*

Proof. Let $\mathcal{M}(\pi_{XZ}) = \{p_{XZ} : p_{XZ}(x, z) \leq \pi_{XZ}(x, z), (x, z) \in \mathbf{X} \times \mathbf{Z}\}$ and $\mathcal{M}(\pi_{YZ}) = \{p_{YZ} : p_{YZ}(y, z) \leq \pi_{YZ}(y, z), (y, z) \in \mathbf{Y} \times \mathbf{Z}\}$. The conditional independence means that any probability distribution $p(x, y, z)$ from $\mathcal{M}(\pi_{XYZ})$ can be expressed as

$$p(x, y, z) = \frac{p_{XZ}(x, z) \cdot p_{YZ}(y, z)}{p_{XZ}^{\downarrow Z}(z)},$$

for $p_{XZ} \in \mathcal{M}(\pi_{XZ})$ and $p_{YZ} \in \mathcal{M}(\pi_{YZ})$ such that $p_{XZ}^{\downarrow Z} = p_{YZ}^{\downarrow Z}$. As $\pi(x, y, z)$ should be the upper envelope of $\mathcal{M}(\pi_{XYZ})$, it can be expressed as

$$\begin{aligned}
\pi(x, y, z) & \\
&= \sup_{p \in \mathcal{M}(\pi_{XYZ})} p(x, y, z) \\
&= \sup_{\substack{p_{XZ} \in \mathcal{M}(\pi_{XZ}) \\ p_{YZ} \in \mathcal{M}(\pi_{YZ})}} \frac{p_{XZ}(x, z) \cdot p_{YZ}(y, z)}{p_{XZ}^{\downarrow Z}(z)} \\
&= \sup_{p_{XZ} \in \mathcal{M}(\pi_{XZ})} \frac{p_{XZ}(x, z)}{p_{XZ}^{\downarrow Z}(z)} \\
&\quad \cdot \sup_{p_{YZ} \in \mathcal{M}(\pi_{YZ})} p_{YZ}(y, z) \\
&= \pi_{X|Z}(x|z) \cdot \pi_{YZ}(y, z).
\end{aligned}$$

In the last equality we utilized the fact that π_{YZ} is the upper envelope of $\mathcal{M}(\pi_{YZ})$, and as proven by Walley and de Cooman [18], the same holds for $\pi_{X|Z}$, if the conditional distribution is obtained by Dempster's conditioning rule, i.e. is based on the product t-norm. ■

The reverse implication is not valid, as can be seen from the following simple example.

Example 2 *(Continued)* One can easily see that the set $\mathcal{M}(\pi_T)$ contains a probability distribution $(.15, .3, .35, .2)$ which is not contained in $\mathcal{M}(\pi_1 \cdot \pi_2)$, as there does not exist a linear combination of its extreme points equal to this distribution. ♢

The implication can be reversed in the totally uninformative case, i.e. when $\pi \equiv 1$. We conjecture that it is the only case.

6 Conclusions

We have presented two different notions of conditional independence in evidence theory proposed by Ben Yaghlane *et al.* and Jiroušek, both of them satisfying semi-graphoid axioms. While the first one is not consistent with marginalization, the latter does not suffer from this drawback.

Nevertheless, we have shown, that although both of them are generalizations of conditional independence concept in probability theory, they cannot be applied to possibility theory.

As a solution of this problem we proposed to use strong independence of credal sets and demonstrated its relation to possibilistic conditional T-independence, more exactly we proved that conditional strong independence implies conditional T-independence based on the product t-norm but not vice versa.

Acknowledgements

The research was partially supported by GA ČR under grant 201/08/0539, GA AV ČR under grant A100750603 and and MŠMT under grant 2C06019.

References

[1] R. G. Almond, *Graphical belief modelling.* Chapman & Hall, 1995.

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

[3] L. M. de Campos, J. F. Huete and S. Moral, Possibilistic independence, in: *Proceedings of EUFIT 95 (Third European Congress on Intelligent Techniques and Soft Computing)*, Verlag Mainz und Wisseschaftverlag, Aachen, 1995, 69–73.

[4] L. M. de Campos and J. F. Huete, Independence concepts in possibility theory: Part 1, *Fuzzy sets and systems* **103** (1999), 127–152.

[5] L. M. de Campos and J. F. Huete, Independence concepts in possibility theory: Part 2, *Fuzzy sets and systems* **103** (1999), 487–505.

[6] 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.

[7] P. Fonck, Conditional independence in possibility theory, in: *Proceedings of 10th Conference on Uncertainty in Artificial Intelligence UAI'94,* eds. R. L. de Mantaras and P. Poole, Morgan Kaufman, San Francisco, 1994, 221–226.

[8] R. Jiroušek, Composition of probability measures on finite spaces, in: *Proceedings of 13th Conference on Uncertainty in Artificial Intelligence UAI'97,* eds. D. Geiger, P. P. Shenoy, Morgan Kaufman, San Francisco, 1997, 274–281.

[9] R. Jiroušek, On a conditional irrelevance relation for belief functions basd on the operator of composition, in: *Dynamics of Knowledge and Belief,*(Ch. Beierle, G. Kern-Isberner, eds.) Proceedings of the Workshop at the 30th Annual German Conference on Artificial Intelligence, Fern Universitt in Hagen, Osnabrck, 2007, pp. 28-41.

[10] R. Jiroušek, J. Vejnarová and M. Daniel, Compositional models for belief functions, *Proceedings of ISIPTA'07,* eds. G. de Cooman, J. Vejnarová, M. Zaffalon, 243–252.

[11] S. Moral, A.Cano, Strong conditional independence for credal sets, *Ann. of Math. and Artif. Intell.*, **35** (2002), 295–321.

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

[13] W. Spohn, Stochastic independence, causal independence and shieldability, *J. Phil. Logic* **9** (1980), 73–99.

[14] M. Studený, Formal properties of conditional independence in different calculi of artificial intelligence, in: *Proceedings of ECSQARU'93,* eds. K. Clarke, R Kruse, S. Moral, Springer-Verlag, pp. 341–348.

[15] J. Vejnarová, Possibilistic independence and operators of composition of possibility measures, in: *Prague Stochastics'98,* eds. M. Hušková, J. Á. Víšek, P. Lachout, JČMF, Prague, 1998, 575–580.

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

[17] J. Vejnarová, Markov properties and factorization of possibility distributions. *Annals of Mathematics and Artificial Intelligence,* **35** (2002), 357–377.

[18] P. Walley and G. de Cooman, Coherence rules for defining conditional possibility, *Int. J. of Approximate Reasoning*, **21** (1999), 63–107.