A Thorough Comparison of Two Conditional Independence Concepts for Belief Functions

Jiřina Vejnarová
Institute of Information Theory and Automation of the ASCR
Pod Vodárenskou věží 4, 182 08 Prague, Czech Republic
Email: vejnar@utia.cas.cz
University of Economics, Prague

***Abstract*—Stochastic conditional independence plays an important role in the application of probability theory into the field of artificial intelligence. From the comparison of complexity of models based on probability distributions and those based on belief functions it is obvious, that it is even more important in the latter framework. In this contribution we compare two conditional independence concepts (conditional non-interactivity and conditional independence) from various points of view. We will concentrate not only to their formal properties, but also to their unconditional versions, their relationship to stochastic conditional independence, number of focal elements of basic assignments satisfying the respective conditional independence constraints, the complexity of their checking, their consistency with marginalization and, naturally, also their mutual relationship.**

Keywords: conditional non-interactivity, conditional independence

I. Introduction

In the application of probabilistic models to the field of artificial intelligence (but not only to it) the most important problem is that it is necessary to model a great number of variables (usually hundreds or even thousands). Therefore, the most frequently used models are so-called *graphical Markov models* (the most popular representatives are Bayesian networks), where the problem of multidimensionality is solved using the notion of conditional independence, which enables factorization of a multidimensional probability distribution into small parts, usually marginal or conditional low-dimensional distributions (or generally into low-dimensional factors). Usually these low-dimensional distributions are the "building blocks" of the resulting multidimensional model.

Such a factorization not only decreases the storage requirements for representation of a multidimensional distribution, but it usually induces efficient computational procedures allowing inference from these models as well. Many results analogous to those concerning conditional independence, Markov properties and factorization from probabilistic framework were also achieved in possibility theory [10], [11].

It is easy to realize that our need of efficient methods for representation of probability and possibility distributions (requiring an exponential number of parameters) logically leads us to greater 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 their representation are superexponential.

The contribution is organized as follows. After a short overview of necessary terminology and notation (Section II), in Section III we recall two concepts of conditional independence. In Section IV we compare them from different points of view: their formal properties, unconditional versions, relationship to stochastic conditional independence, number of focal elements of a basic assignments satisfying the respective conditional independence constraints, the complexity of its checking, their consistency with marginalization and, finally, their mutual relationship.

II. Basic notions

A. Set projections and joins

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 *multidimensional 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 two opposite operations, which will be called a cylindrical extension and a join.

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

A *cylindrical extension* $A^{\uparrow K}$ of a set $A \subseteq \mathbf{X}_M$ to $\mathbf{X}_K (M \subseteq K)$ is defined by the equality

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

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 their join 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,

$$A \bowtie B = A^{\uparrow K \cup L} \cap B^{\uparrow K \cup L}. \tag{1}$$

B. Set functions

In evidence theory [7] (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 $m(\emptyset) = 0$.

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

$$\begin{aligned} Bel(A) &= \sum_{B \subseteq A} m(B), \\ Pl(A) &= \sum_{B \cap A \neq \emptyset} m(B), \end{aligned}$$

respectively.

In addition to belief and plausibility measures, *commonality function* can also be obtained from basic assignment m:

$$Q(A) = \sum_{B \supseteq A} m(B).$$

The last notion plays an important role in the definition of so-called conditional non-interactivity of variables (cf. Section III-A) and in Shenoy's valuation-based systems [8] — commonality functions are a special type of proper normal valuations.

A set $A \in \mathcal{P}(\mathbf{X}_N)$ is a *focal element* if $m(A) > 0$. A pair $(\mathcal{F}, m)$, where $\mathcal{F}$ is the set of all focal elements, is called a *body of evidence*. A basic assignment is called *Bayesian* if all its focal elements are singletons.

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

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

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

Let us note that $m^{\downarrow \emptyset} \equiv 1$ for arbitrary basic assignment m.

Analogously, $Q^{\downarrow M}$ will denote the corresponding marginal commonality function.

Having two basic assignments m_1 and m_2 on $\mathbf{X}_K$ and $\mathbf{X}_L$, respectively $(K, L \subseteq N)$, 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.

According to the above convention m_1 and m_2 are projective whenever $K \cap L = \emptyset$.

Dempster's rule of combination has been from its appearance frequently criticized by many authors. Therefore, many alternatives to it were suggested by various authors. From the viewpoint of this paper, the most important is *conjunctive rule*, which is, in fact, unnormalized Dempster's rule defined for any C by the formula

$$(m_1 \circledcirc m_2)(C) = \sum_{A, B \subseteq \mathbf{X}_K A \cap B = C} m_1(A) \cdot m_2(B).$$

It can easily be generalized to the case when m_1 is defined on X_K and m_2 is defined on X_L $(K \neq L)$ in the following way:

$$(m_1 \circledcirc m_2)(C) = \sum_{\substack{A \subseteq \mathbf{X}_K, B \subseteq \mathbf{X}_L \\ A^{\uparrow L \cup K} \cap B^{\uparrow L \cup K} = C}} m_1(A) \cdot m_2(B). \tag{2}$$

for any $C \in \mathbf{X}_{K \cup L}$.

III. Two concepts of conditional independence

A. Conditional non-interactivity

Ben Yaghlane et al. [2] introduced the notion of conditional non-interactivity only for the case of three variables. In order to unify the notation throughout the paper we rewrite it in the following way.

Definition 1. Let m be a basic assignment on $\mathbf{X}_N$ and $K, L, M \subset N$ be disjoint, $K \neq \emptyset \neq L$. Groups of variables X_K and X_L are *conditionally non-interactive given* X_M *with respect to* m, denoted by $X_K \perp_m X_L | X_M$ if and only if the equality

$$m^{\downarrow K \cup L \cup M} \circledcirc m^{\downarrow M}(A) = m^{\downarrow K \cup M} \circledcirc m^{\downarrow L \cup M}(A) \tag{3}$$

is satisfied for any $A \subseteq \mathbf{X}_{K \cup L \cup M}$.

Furthermore, they proved that (3) holds if and only if the equality

$$\begin{aligned} &Q^{\downarrow K \cup L \cup M}(A) \cdot Q^{\downarrow M}(A^{\downarrow M}) \\ &\quad = Q^{\downarrow K \cup M}(A^{\downarrow K \cup M}) \cdot Q^{\downarrow L \cup M}(A^{\downarrow L \cup M}) \end{aligned} \tag{4}$$

holds for any $A \subseteq \mathbf{X}_{K\cup L\cup M}$.

Let us note that (4) is more common expression of conditional non-interactivity than (3) and it is a special case of the definition of conditional independence in valuation-based systems introduced by Shenoy [8] (for valuations expressed by means of commonality functions). Nevertheless, in valuation-based systems commonality function is a primitive concept and basic assignment is derived by formula

$$m(A) = \sum_{B\supseteq A} (-1)^{|B\setminus A|} Q(B).$$

Nevertheless, this notion of independence does not seem to be appropriate for construction of multidimensional models. As already mentioned by Studený [9], it is not consistent with marginalization. What that means will be shown in Subsection IV-F.

Therefore, instead of the conditional non-interactivity, in [4], [12] we proposed to use another notion of conditional independence recalled in the following subsection.

B. Conditional independence

Let us start this section by recalling the notion of random sets independence [3]:[3] 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 (and denote it by $K \perp\!\!\!\perp L\ [m]$) 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.

In [12] we generalized this notion in the following way.

Definition 2. 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

$$\begin{aligned} & m^{\downarrow K\cup L\cup M}(A) \cdot m^{\downarrow M}(A^{\downarrow M}) \\ &= \ m^{\downarrow K\cup M}(A^{\downarrow K\cup M}) \cdot m^{\downarrow L\cup M}(A^{\downarrow L\cup M}) \end{aligned} \quad (5)$$

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.

Let us note that (5) resembles, from the formal point of view, the definition of stochastic conditional independence [6]. Aparently (5) and (4) are almost identical except that (5) uses m instead of Q. But it is not true, as (4) must be satisfied for *any* $A \subseteq \mathbf{X}_{K\cup L\cup M}$, while (5) only for those sets which can be expressed as joins of their marginals.

In the following section we will present a thorough comparison of these two concepts from various points of view.

[3]Klir [5] uses the notion *non-interactivity*.

IV. Comparison

A. Marginal case

As already mentioned in Section III, conditional independence is a generalization of random set independence and similar relationship holds for conditional and unconditional non-interactivity (cf. [2]).

Let us present an assertion showing that conditional non-interactivity and conditional independence are identical if the condition is empty.

Lemma 1. *Let* K, L *be disjoint, then* $K \perp\!\!\!\perp L\ [m]$ *if and only if*

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

for all $A \subseteq \mathbf{X}_{K\cup L}$.

Proof can be found in [4].

So, it can be concluded that conditional non-interactivity and conditional independence are (different) generalizations of the same concept. These differences will be in the center of our attention in the rest of this section.

B. Formal properties

Among the properties satisfied by the ternary relation $K \perp\!\!\!\perp L|M\ [m]$, the following are of principal importance:

(A1) $K \perp\!\!\!\perp L|M\ [m] \ \Rightarrow \ L \perp\!\!\!\perp K|M\ [m]$,
(A2) $K \perp\!\!\!\perp L\cup M|I\ [m] \ \Rightarrow \ K \perp\!\!\!\perp M|I\ [m]$,
(A3) $K \perp\!\!\!\perp L\cup M|I\ [m] \ \Rightarrow \ K \perp\!\!\!\perp L|M\cup I\ [m]$,
(A4) $K \perp\!\!\!\perp L|M\cup I\ [m] \ \wedge \ K \perp\!\!\!\perp M|I\ [m] \ \Longrightarrow \ K \perp\!\!\!\perp L\cup M|I\ [m]$,
(A5) $K \perp\!\!\!\perp L|M\cup I\ [m] \ \wedge \ K \perp\!\!\!\perp M|L\cup I\ [m] \ \Longrightarrow \ K \perp\!\!\!\perp L\cup M|I\ [m]$.

Let us recall that stochastic conditional independence satisfies the so-called *semigraphoid* properties (A1)–(A4) for any probability distribution, while axiom (A5) is satisfied only for strictly positive probability distributions. Analogous results were proven in [12] also for conditional independence presented in Definition 2.

Theorem 1. *Conditional independence satisfies (A1)–(A4).*

Theorem 2. *Let* m *be a basic assignment on* $\mathbf{X}_N$ *such that* $m(A) > 0$ *if and only if* $A = \times_{i\in N} A_i$*, where* A_i *is a focal element on* $\mathbf{X}_i$*. Then (A5) is satisfied.*

Conditional non-interactivity referred to in Section III-A, on the other hand, satisfies axioms (A1)–(A5) for general basic assignment m, as stated in [2].

C. Bayesian case

Now let us present, that both concepts are equivalent in the case of Bayesian basic assignment.

Lemma 2. *Let* m *be a Bayesian basic assignment on* $\mathbf{X}_N$*. Then*

$$X_K \perp_m X_L | X_M \Longleftrightarrow K \perp\!\!\!\perp L|M\ [m]$$

for any three disjoint subsets K, L, M ($K, L \neq \emptyset$) of N.

Proof. To prove the equivalence it is enough to realize, that the only focal elements of a Bayesian basic assignment are singletons, and therefore $m(A) = Q(A)$ for any $A \subseteq \mathbf{X}_N$. Furthermore, as $m(A) = 0$ for all non-singletons, then it is obvious that also $m(A) = 0$ for $A^{\downarrow K\cup L\cup M} \neq A^{\downarrow K\cup M} \bowtie A^{\downarrow L\cup M}$.

In this case both (4) and (5) become

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

for all $x \in \mathbf{X}_N$, otherwise $m(A) = 0$ and both equalities are trivially satisfied. □

From the proof of Lemma 2 one can immediately see that both concepts generalize the well-known concept of stochastic conditional independence. From this viewpoint the statement that conditional non-interactivity satisfies so-called graphoid axioms seems to be questionable.

D. Focal elements

From the definition of $K \perp\!\!\!\perp L|M\ [m]$ it is obvious that the only focal elements under this property are those, which can be expressed as the joins of their marginals.

It has been proven in [2] that if $X_K \perp_m X_L|X_M$, then the focal elements of m belong to the set of X_M-layered rectangles.[4]

From (1) and from the fact that any intersection of cylindrical extensions is X_M-layered rectangle (cf. [2]) it follows that number of focal elements under conditional independence is not bigger than that under conditional non-interactivity.

E. Complexity

From the definition of conditional non-interactivity (or from its equivalent characterization) it is obvious that to test one conditional non-interactivity statement $X_K \perp_m X_L|X_M$, it is necessary check validity of $2^{|\mathbf{X}_{K\cup L\cup M}|} - 1$ equalities (as the last one is trivial).

From the definition of conditional independence, on the other hand, it is evident that it is enough to check validity of a smaller number of equalities: only for those subsets which can be expressed as joins of their marginals, as by Definition 2 for remaining sets $m(A) = 0$. Upper bound of the number of equalities to be checked is contained in the following lemma.

Lemma 3. *The number of equalities to be checked in order to test $K \perp\!\!\!\perp L|M\ [m]$ is smaller than $2^{|\mathbf{X}_M|\cdot(|\mathbf{X}_K|+|\mathbf{X}_L|)}$.*

Proof. The only focal elements of m such that $K \perp\!\!\!\perp L|M\ [m]$ are such that $A = A^{\downarrow K\cup M} \bowtie A^{\downarrow L\cup M}$, i.e. the number of focal elements must be smaller than the product of numbers of the focal elements of $m^{\downarrow K\cup M}$ and $m^{\downarrow L\cup M}$, i.e. smaller than $2^{|\mathbf{X}_{K\cup M}|} \cdot 2^{|\mathbf{X}_{K\cup L}|}$. □

This upper bound is very rough, as can be seen from the following simple example.

[4]Let us remind that X_M-layered rectangle is a set in $\mathbf{X}_{K\cup L\cup M}$ which is for nay fixed value $x_M \in \mathbf{X}_M$ a rectangle in $\mathbf{X}_K \times \mathbf{X}_L$.

Example 1. Let X_1, X_2 and X_3 be three binary variables with values in $\mathbf{X}_1 = \{a_1, \bar{a}_1\}, \mathbf{X}_2 = \{a_2, \bar{a}_2\}$, $\mathbf{X}_3 = \{a_3, \bar{a}_3\}$ described by a basic assignment m. Let us assume that we have to check the relation $1 \perp\!\!\!\perp 2|3\ [m]$. The potential focal elements are $A \subseteq \mathbf{X}_1 \times \mathbf{X}_2 \times \mathbf{X}_3$ such that $A = A^{\downarrow 13} \bowtie A^{\downarrow 23}$ and

- either $A^{\downarrow i3} \in \mathcal{P}(\{(a_i, a_3), (\bar{a}_i, a_3)\}) \setminus \emptyset$,
- or $A^{\downarrow i3} \in \mathcal{P}(\{(a_i, \bar{a}_3), (\bar{a}_i, \bar{a}_3)\}) \setminus \emptyset$,
- or $C^{\downarrow i3} \in \mathcal{P}(\{(a_i, a_3), (\bar{a}_i, a_3), (a_i, \bar{a}_3), (\bar{a}_i, \bar{a}_3)\}) \setminus\mathcal{P}(\{(a_i, a_3), (\bar{a}_i, a_3)\}) \setminus \mathcal{P}(\{(a_i, \bar{a}_3), (\bar{a}_i, \bar{a}_3)\})$.

We have only 9 sets in both the first and second cases and 81 sets in the third case, i.e. only 99 equalities of the form (5) must be checked (and not 256, as Lemma 3 says). ♢

It is evident, even from the rough upper bound presented in Lemma 3, that, in the general case, to test $K \perp\!\!\!\perp L|M\ [m]$ is computationally less demanding than to test $X_K \perp_m X_L|X_M$ as

$$2^{|\mathbf{X}_M|\cdot(|\mathbf{X}_K|+|\mathbf{X}_L|)} \leq 2^{|\mathbf{X}_M|\cdot|\mathbf{X}_K|\cdot|\mathbf{X}_L|}$$

if both $\mathbf{X}_K$ and $\mathbf{X}_L$ contain at least two elements.

F. Consistency with marginalization

An independence concept is *consistent with marginalization* [9] iff for arbitrary projective basic assignments (probability distributions, possibility distributions, etc.) m_1 on $\mathbf{X}_K$ and m_2 on $\mathbf{X}_L$ there exists a basic assignment (probability distribution, possibility distribution, etc.) on $\mathbf{X}_{K\cup L}$ satisfying this independence concept and having m_1 and m_2 as its marginals.

In [2] one can find the following example (originally suggested by Studený) showing that application of (4) to two projective basic assignments may lead to a model which is beyond the framework of evidence theory.

Example 2. Let X_1, X_2 and X_3 be three binary variables with values in $\mathbf{X}_1 = \{a_1, \bar{a}_1\}, \mathbf{X}_2 = \{a_2, \bar{a}_2\}$, $\mathbf{X}_3 = \{a_3, \bar{a}_3\}$ and m_1 and m_2 be two basic assignments on $\mathbf{X}_1 \times \mathbf{X}_3$ and $\mathbf{X}_2 \times \mathbf{X}_3$ respectively, both of them having only two focal elements:

$$\begin{aligned} m_1(\{(a_1, \bar{a}_3), (\bar{a}_1, \bar{a}_3)\}) &= .5, \\ m_1(\{(a_1, \bar{a}_3), (\bar{a}_1, a_3)\}) &= .5, \\ m_2(\{(a_2, \bar{a}_3), (\bar{a}_2, \bar{a}_3)\}) &= .5, \\ m_2(\{(a_2, \bar{a}_3), (\bar{a}_2, a_3)\}) &= .5. \end{aligned} \tag{6}$$

Since their marginals are projective

$$m_1^{\downarrow 3}(\{\bar{a}_3\}) = m_2^{\downarrow 3}(\{\bar{a}_3\}) = .5,$$
$$m_1^{\downarrow 3}(\{a_3, \bar{a}_3\}) = m_2^{\downarrow 3}(\{a_3, \bar{a}_3\}) = .5,$$

there exists (at least one) common extension of both of them, but none of them is such that it would imply conditional non-interactivity of X_1 and X_2 given X_3. Namely, the application of equality (4) to basic assignments m_1 and m_2 leads to the following values of the joint "basic assignment":

$$\begin{aligned}\bar{m}(\mathbf{X}_1 \times \mathbf{X}_2 \times \{\bar{a}_3\}) &= .25,\\ \bar{m}(\mathbf{X}_1 \times \{a_2\} \times \{\bar{a}_3\}) &= .25,\\ \bar{m}(\{a_1\} \times \mathbf{X}_2 \times \{\bar{a}_3\}) &= .25,\\ \bar{m}(\{(a_1, a_2, \bar{a}_3), (\bar{a}_1, \bar{a}_2, a_3)\}) &= .5,\\ \bar{m}(\{(a_1, a_2, \bar{a}_3)\}) &= -.25,\end{aligned}$$

which is outside of evidence theory. ♢

This problem is solved in [2] by the application of (3) to the marginals. The result is now basic assignment, but it does not keep the original marginals, as can be seen from the following example.

Example 2. *(Continued)* Application of (3) to marginals (6) leads to the joint basic assignment

$$\begin{aligned}\hat{m}(\mathbf{X}_1 \times \mathbf{X}_2 \times \{\bar{a}_3\}) &= .25,\\ \hat{m}(\mathbf{X}_1 \times \{a_2\} \times \{\bar{a}_3\}) &= .25,\\ \hat{m}(\{a_1\} \times \mathbf{X}_2 \times \{\bar{a}_3\}) &= .25,\\ \hat{m}(\{(a_1, a_2, \bar{a}_3), (\bar{a}_1, \bar{a}_2, a_3)\}) &= .25\end{aligned} \tag{7}$$

with marginal basic assignments

$$\begin{aligned}m^{\downarrow 13}(\{(a_1, \bar{a}_3), (\bar{a}_1, \bar{a}_3)\}) &= .5,\\ m^{\downarrow 13}(\{(a_1, \bar{a}_3)\}) &= .25,\\ m^{\downarrow 13}(\{(a_1, \bar{a}_3), (\bar{a}_1, a_3)\}) &= .25,\\ m^{\downarrow 23}(\{(a_2, \bar{a}_3), (\bar{a}_2, \bar{a}_3)\}) &= .5,\\ m^{\downarrow 23}(\{(a_2, \bar{a}_3)\}) &= .25,\\ m^{\downarrow 23}(\{(a_2, \bar{a}_3), (\bar{a}_2, a_3)\}) &= .25.\end{aligned}$$

♢

These different results obtained by the application of one conditional non-interactivity concept seem to be, at the first glance, surprising. Nevertheless, it is caused by the fact, that (as already said above), *there does not exist any joint basic assignment keeping the prescribed marginals such that it would imply conditional non-interactivity of X_1 and X_2 given* X_3. Therefore, if we try to find it, something gets wrong: when (4) is applied, we obtain "basic assignment" which is not nonnegative for all subsets; application of (3), on the other hand, ensures, that the result is basic assignment, but simultaneously it "spoils" the prescribed marginals.

Therefore, this solution is not completely satisfactory — although the resulting model belongs to the evidence theory framework, the marginals are different (i.e. the original information has been changed).

The following assertion expresses the fact (already mentioned above) that the concept of conditional independence $K \perp\!\!\!\perp L|M\ [m]$ is consistent with marginalization. Moreover, it presents a form expressing the joint basic assignment by means of its marginals.

Theorem 3. *Let m_1 and m_2 be projective basic assignments on $\mathbf{X}_K$ and $\mathbf{X}_L$, respectively. Let us define a basic assignment m on $\mathbf{X}_{K\cup L}$ by the formula*

$$m(A) = \frac{m_1(A^{\downarrow K}) \cdot m_2(A^{\downarrow L})}{m_2^{\downarrow K\cap L}(A^{\downarrow K\cap L})} \tag{8}$$

for $A = A^{\downarrow K} \bowtie A^{\downarrow L}$ such that $m_1^{\downarrow K\cap L}(A^{\downarrow K\cap L}) > 0$ and $m(A) = 0$ otherwise. Then

$$\begin{aligned}m^{\downarrow K}(B) &= m_1(B),\\ m^{\downarrow L}(C) &= m_2(C)\end{aligned}$$

for any $B \in \mathbf{X}_K$ and $C \in \mathbf{X}_L$, respectively, and $(K \setminus L) \perp\!\!\!\perp (L \setminus K)|(K \cap L)\ [m]$. Furthermore, m is the only basic assignment possessing these properties.

Proof can be found in [12].

Let us close this section by demonstrating application of the conditional independence notion (and Theorem 3) to Example 2.

Example 2. *(Continued)* Let us go back to the problem of finding a common extension of basic assignments m_1 and m_2 defined by (6). Theorem 3 says that for basic assignment m

$$\begin{aligned}m(\mathbf{X}_1 \times \mathbf{X}_2 \times \{\bar{a}_3\}) &= .5,\\ m(\{(a_1, a_2, \bar{a}_3), (\bar{a}_1, \bar{a}_2, a_3)\}) &= .5,\end{aligned} \tag{9}$$

obtained from (6) by (8), variables X_1 and X_2 are conditionally independent given X_3. ♢

G. Mutual relationship

Conditional independence does not imply conditional non-interactivity, as can be seen from the following example.

Example 3. Let X_1, X_2 and X_3 be three binary variables with the joint basic assignment defined by (9). As can be seen from the last part of Example 2 variables X_1 and X_2 are conditionally independent given X_3, but they are not conditionally non-interactive, as e.g.

$$\begin{aligned}&Q(\mathbf{X}_1 \times \mathbf{X}_2 \times \{\bar{a}_3\}) \cdot Q^{\downarrow 3}(\{\bar{a}_3\})\\ &\quad = 0.5 \times 1 \neq 0.5 \times 0.5\\ &\quad = Q^{\downarrow 13}(\mathbf{X}_1 \times \{\bar{a}_3\}) \cdot Q^{\downarrow 23}(\mathbf{X}_2 \times \{\bar{a}_3\}).\end{aligned}$$

♢

On the other hand, neither conditional non-interactivity implies conditional independence. It is demonstrated by the following example.

Example 4. Let X_1, X_2 and X_3 be three binary variables with the joint basic assignment defined by (7). As stated in [2] X_1 and X_2 are conditionally non-interactive given X_3, but they are not conditionally independent, as e.g.

$$\begin{aligned}&m(\mathbf{X}_1 \times \mathbf{X}_2 \times \{\bar{a}_3\}) \cdot m^{\downarrow 3}(\{\bar{a}_3\})\\ &\quad = 0.25 \times 0.75 \neq 0.5 \times 0.5\\ &\quad = m^{\downarrow 13}(\mathbf{X}_1 \times \{\bar{a}_3\}) \cdot m^{\downarrow 23}(\mathbf{X}_2 \times \{\bar{a}_3\}).\end{aligned}$$

♢

From these two simple examples one can deduce that none of these two conditional independence concepts is stronger than the other.

V. Conclusions

We recalled two conditional independence concepts in evidence theory, compared them from various points of view and realized:

- They are equivalent to each other if the condition is empty, i.e. unconditional non-interactivity coincides with unconditional random set independence.
- They are equivalent to each other for Bayesian basic assignments. In this case they collapse to stochastic conditional independence.
- From the viewpoint of formal properties conditional non-interactivity seems to be preferable, as it satisfies (as stated in [2]) so-called graphoid properties, while conditional independence only semigraphoid ones. This superiority is somewhat relativized by the preceding statement (as stochastic conditional independence satisfies only semigraphoid properties).
- The number of (potential) focal elements under conditional independence is not greater than that under conditional non-interactivity, as a join of its projections is a special case of X_M-layered rectangle.
- The complexity of checking conditional non-interactivity is substantially higher, as it is necessary to check prescribed equality for all subsets of the frame of discernment in question, while for conditional independence only for those sets, which are joins of its projections.
- From the multidimensional model construction point of view the most substantial difference is that conditional non-interactivity is not consistent with marginalization, while conditional independence is.
- None of these concepts is stronger that the other one.

It can be summarized that the only disadvantage of conditional independence in comparison with conditional non-interactivity is, that it does not generally satisfy axiom (A5), while we can see two drawbacks of conditional non-interactivity: complexity of checking the conditional independence statements and mainly the fact that it is not consistent with marginalization, which leads to multidimensional models not keeping the original information.

Acknowledgement

The work of the author was supported by the grant GA ČR 201/09/1891, by the grant GA AV ČR A100750603 and by the grant MŠMT 2C06019.

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), pp. 479–513

[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] 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, Ghent, June 29 – July 2, 1999, pp. 121–130.

[4] R. Jiroušek, J. Vejnarová, "Compositional models and conditional independence in Evidence Theory," accepted to *International Journal of Approximate Reasoning*.

[5] G. J. Klir, *Uncertainty and Information. Foundations of Generalized Information Theory.* Location: Wiley, Hoboken, 2006.

[6] S. L. Lauritzen, *Graphical Models*. Location: Oxford University Press, 1996.

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

[8] P. P. Shenoy, "Conditional independence in valuation-based systems," *Int. J. Approx. reasoning*, vol. 10, pp. 203–234, 1994.

[9] M. Studený, "Formal properties of conditional independence in different calculi of artificial intelligence." *Proceedings of European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty ECSQARU'93*, eds. K. Clarke, R. Kruse, S. Moral, Springer-Verlag, 1993, pp. 341–348.

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

[11] J. Vejnarová, "Markov properties and factorization of possibility distributions." *Annals of Mathematics and Artificial Intelligence,* vol. 35, no. 1–4, pp. 357–377, 2002.

[12] J. Vejnarová, "On conditional independence in evidence theory," In: Proceedings of the 6th International Symposium on Imprecise Probability: Theories and Applications, ISIPTA'09. Durham, UK, 2009, pp. 431–440.