A Note on Local Computations in Dempster-Shafer Theory of Evidence

Radim Jiroušek
Faculty of Management of University of Economics, and
Institute of Information Theory and Automation, Academy of Sciences
Czech Republic
radim@utia.cas.cz

Abstract

When applying any technique of multidimensional models to problems of practice, one always has to cope with two problems: it is necessary to have a possibility to represent the models with a "reasonable" number of parameters and to have sufficiently efficient computational procedures at one's disposal. When considering graphical Markov models in probability theory, both of these conditions are fulfilled; various computational procedures for decomposable models are based on the ideas of local computations, whose theoretical foundations were laid by Lauritzen and Spiegelhalter.

The presented contribution studies a possibility of transferring these ideas from probability theory into Dempster-Shafer theory of evidence. The paper recalls decomposable models, discusses connection of the model structure with the corresponding system of conditional independence relations, and shows that under special additional conditions, one can locally compute specific basic assignments which can be considered to be conditional.

Keywords. Multidimensional models, graphical models, conditional independence, factorisation, computations.

1 Introduction

The great advantage of Dempster-Shafer theory [5, 18] is the fact that it generalises classical probability theory in the way that one can easily describe not only uncertainty but also vagueness (ignorance). Nevertheless, the disadvantage of this approach stems from the fact that belief functions cannot be represented by a point function (like density in probability theory); instead, one has to manipulate with set functions, which leads to exponential increase of algorithmic complexity of all the necessary computational procedures.

With regard to probability theory, substantial decrease of computational complexity was achieved with the help of Graphical Markov Models (GMM), a technique developed in the last quarter of the last century. Here we specifically have in mind a technique based on local computations for which theoretical background was laid by Lauritzen and Spiegelhalter [17]. Its basic idea can be expressed in a few words: a multidimensional distribution represented by a Bayesian network is first converted into a decomposable model, which allows for efficient computation of conditional probabilities.

Studying properly probabilistic GMM one can realise that it is a notion of *conditional independence* (which is closely connected with a notion of *factorisation*) that makes it possible to represent multidimensional probability distributions efficiently. A goal of this paper is to make a brief survey summarising results concerning decomposable models within Dempster-Shafer theory of evidence presented in [10, 11, 12]. In addition to this we will show that, even in Dempster-Shafer theory, one can employ the basic ideas of Lauritzen and Spiegelhalter and compute "conditional" basic assignments locally.

1.1 Notation

In this paper we consider a finite multidimensional space $\mathbf{X}_N = \mathbf{X}_1 \times \mathbf{X}_2 \times \ldots \times \mathbf{X}_n$, and its subspaces (for all $K \subseteq N$)

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

For a point $x = (x_1, x_2, \ldots, x_n) \in \mathbf{X}_N$ its projection into subspace $\mathbf{X}_K$ is denoted $x^{\downarrow K} = (x_{i,i \in K})$, and for $A \subseteq \mathbf{X}_N$

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

By a *join* of two sets $A \subseteq \mathbf{X}_K$ and $B \subseteq \mathbf{X}_L$ we 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 $A \otimes B = A \times B$, if $K = L$ then $A \otimes B = A \cap B$.

In view of this paper it is important to realise that if $x \in C \subseteq \mathbf{X}_{K \cup L}$, then $x^{\downarrow K} \in C^{\downarrow K}$ and $x^{\downarrow L} \in C^{\downarrow L}$, which means that always $C \subseteq C^{\downarrow K} \otimes C^{\downarrow L}$. However, it does not mean that $C = C^{\downarrow K} \otimes C^{\downarrow L}$. For example, considering two-dimensional frame of discernment $\mathbf{X}_{\{1,2\}}$ with $\mathbf{X}_i = \{a_i, \bar{a}_i\}$ for both $i = 1, 2$, and $C = \{a_1 a_2, \bar{a}_1 a_2, a_1 \bar{a}_2\}$, one gets

$$\begin{aligned} C^{\downarrow\{1\}} \otimes C^{\downarrow\{2\}} &= \{a_1, \bar{a}_1\} \otimes \{a_2, \bar{a}_2\} \\ &= \{a_1 a_2, \bar{a}_1 a_2, a_1 \bar{a}_2, \bar{a}_1 \bar{a}_2\} \supsetneq C. \end{aligned}$$

1.2 Basic assignments

The role played by a probability distribution in probability theory is replaced by that of a set function in Dempster-Shafer theory: belief function, plausibility function or basic (*probability or belief*) assignment. Knowing one of them, one can derive the remaining two. In this paper we will use almost exclusively basic assignments.

A *basic assignment* m on $\mathbf{X}_K$ $(K \subseteq N)$ is a function

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

for which

$$\sum_{\emptyset \neq A \subseteq \mathbf{X}_K} m(A) = 1.$$

If $m(A) > 0$, then A is said to be a *focal element* of m. Recall that

$$Bel(A) = \sum_{\emptyset \neq B \subseteq A} m(B),$$

and

$$Pl(A) = \sum_{B \subseteq \mathbf{X}_K : B \cap A \neq \emptyset} m(B).$$

Having a basic assignment m on $\mathbf{X}_K$ one can consider its *marginal assignment* on $\mathbf{X}_L$ (for $L \subseteq K$), which is defined (for each $\emptyset \neq B \subseteq \mathbf{X}_L$):

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

1.3 Operator of composition

Compositional models were introduced for probability theory in [8] as an alternative to Bayesian networks for efficient representation of multidimensional measures. They were based on recurrent application of an operator of composition. An analogous operator within the framework of Dempster-Shafer theory was introduced in [14]).

Definition 1 ***Operator of Composition.*** *For two arbitrary basic assignments m_1 on $\mathbf{X}_K$ and m_2 on $\mathbf{X}_L$ $(K \neq \emptyset \neq L)$, a* composition *$m_1 \triangleright m_2$ is defined for each $C \subseteq \mathbf{X}_{K \cup L}$ 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*

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

[c] *in all other cases $(m_1 \triangleright m_2)(C) = 0$.*

Remark 1 First of all, we want to stress that the operator of composition is something other than the famous Dempster's rule of combination [5], or its non-normalised version, the so called *conjunctive combination rule* [1]

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

For example, the operation of composition is (in contrast with the above-mentioned conjunctive combination rule) neither commutative nor associative. While Dempster's rule of combination was designed to combine different (independent) sources of information (it realises fusion of sources), the operator of composition primarily serves for composing pieces of local information (usually coming from one source) into a global model. The notion of composition is therefore closely connected with the notion of *factorisation*. This fact manifests also in the following difference: while for computation of $(m_1 \triangleright m_2)(C)$ it is enough to know only m_1 and m_2 just for the respective projections of set C, computing $(m_1 \circledcirc m_2)(C)$ requires knowledge of, roughly speaking, the entire basic assignments m_1 and m_2.

For further intuitive justification of the operator of composition the reader is referred to [14], where a number of its properties were proved. In view of the forthcoming text, those presented in the following assertion are the most important.

Proposition 1 ***Basic Properties.*** *Let m_1 and m_2 be basic assignments defined on $\mathbf{X}_K, \mathbf{X}_L$, respectively. Then:*

1. *$m_1 \triangleright m_2$ is a basic assignment on $\mathbf{X}_{K \cup L}$;*
2. *$(m_1 \triangleright m_2)^{\downarrow K} = m_1$;*
3. *$m_1 \triangleright m_2 = m_2 \triangleright m_1 \iff m_1^{\downarrow K \cap L} = m_2^{\downarrow K \cap L}$.*

The reader probably noticed that Property 2 guarantees idempotency of the operator and gives a hint about how to get a counterexample to its commutativity. From point 1, one immediately gets that for basic assignments $m_1, m_2, \ldots, m_r$ defined on $\mathbf{X}_{K_1}, \mathbf{X}_{K_2}, \ldots, \mathbf{X}_{K_r}$, respectively, the formula $m_1 \triangleright m_2 \triangleright \ldots \triangleright m_r$ defines a (possibly multidimensional) basic assignment defined on $\mathbf{X}_{K_1 \cup \ldots \cup K_r}$.

2 Controlled associativity

As already mentioned above, the operator of composition is not associative. This means that in fact we do not know what the formula $m_1 \triangleright m_2 \triangleright \ldots \triangleright m_r$ means. To avoid the necessity of using too many parentheses, let us make the following convention. In the formulae like $m_1 \triangleright m_2 \triangleright \ldots \triangleright m_r$, when the order of application of the operators of composition is not controlled by parentheses, the operators will be applied from left to right, i.e.,

$$m_1 \triangleright m_2 \triangleright \ldots \triangleright m_r = (\ldots(m_1 \triangleright m_2) \triangleright \ldots \triangleright m_{r-1}) \triangleright m_r.$$

Nevertheless, when designing a process of local computations for compositional models in D-S theory (which is intended to be an analogy to the process proposed by Lauritzen and Spiegelhalter in [17]), one needs a type of associativity expressed in the following assertion.

Proposition 2 ***Controlled associativity.*** *Let m_1, m_2 and m_3 be basic assignments on $\mathbf{X}_{K_1}, \mathbf{X}_{K_2}$ and $\mathbf{X}_{K_3}$, respectively, such that $K_2 \supseteq K_1 \cap K_3$, and*

$$m_1^{\downarrow K_1 \cap K_2}(C^{\downarrow K_1 \cap K_2}) > 0 \Longrightarrow m_2^{\downarrow K_1 \cap K_2}(C^{\downarrow K_1 \cap K_2}) > 0.$$

Then

$$(m_1 \triangleright m_2) \triangleright m_3 = m_1 \triangleright (m_2 \triangleright m_3).$$

Proof. The goal is to prove that for any $C \subseteq \mathbf{X}_{K_1 \cup K_2 \cup K_3}$

$$((m_1 \triangleright m_2) \triangleright m_3)(C) = (m_1 \triangleright (m_2 \triangleright m_3))(C). \quad (1)$$

We have to distinguish five special cases.

A. $C \neq C^{\downarrow K_1} \otimes C^{\downarrow K_2} \otimes C^{\downarrow K_3}$.
This is the simplest situation because, due to associativity of join,

$$(C^{\downarrow K_1} \otimes C^{\downarrow K_2}) \otimes C^{\downarrow K_3} = C^{\downarrow K_1} \otimes (C^{\downarrow K_2} \otimes C^{\downarrow K_3})$$

and therefore in this case both sides of formula (1) equal 0, which follows from Definition 1 (case [c]).

B. $C = C^{\downarrow K_1} \otimes C^{\downarrow K_2} \otimes C^{\downarrow K_3}$
& $m_2^{\downarrow K_1 \cap K_2}(C^{\downarrow K_1 \cap K_2}) > 0, m_3^{\downarrow K_2 \cap K_3}(C^{\downarrow K_2 \cap K_3}) > 0.$

In this case, under the given assumptions,

$$K_3 \cap (K_1 \cup K_2) = K_3 \cap K_2$$

and therefore

$$((m_1 \triangleright m_2) \triangleright m_3)(C) = \frac{m_1(C^{\downarrow K_1}) \cdot m_2(C^{\downarrow K_2})}{m_2^{\downarrow K_2 \cap K_1}(C^{\downarrow K_2 \cap K_1})} \cdot \frac{m_3(C^{\downarrow K_3})}{m_3^{\downarrow K_3 \cap K_2}(C^{\downarrow K_3 \cap K_2})}.$$

Analogously, we can make the following computations (in the last modification we use the fact that in the considered case $K_1 \cap K_2 \cap K_3 = K_1 \cap K_3$):

$$\begin{aligned}
&(m_1 \triangleright (m_2 \triangleright m_3))(C) \\
&= \frac{m_1(C^{\downarrow K_1}) \cdot (m_2 \triangleright m_3)(C^{\downarrow K_2 \cup K_3})}{(m_2 \triangleright m_3)^{\downarrow K_1 \cap (K_2 \cup K_3)}(C^{\downarrow K_1 \cap (K_2 \cup K_3)})} \\
&= \frac{m_1(C^{\downarrow K_1})}{(m_2 \triangleright m_3)^{\downarrow K_1 \cap (K_2 \cup K_3)}(C^{\downarrow K_1 \cap (K_2 \cup K_3)})} \cdot \frac{m_2(C^{\downarrow K_2}) \cdot m_3(C^{\downarrow K_3})}{m_3^{\downarrow K_2 \cap K_3}(C^{\downarrow K_2 \cap K_3})} \\
&= \frac{m_1(C^{\downarrow K_1}) \cdot m_3^{\downarrow K_1 \cap K_2 \cap K_3}(C^{\downarrow K_1 \cap K_2 \cap K_3})}{m_2^{\downarrow K_1 \cap K_2}(C^{\downarrow K_1 \cap K_2}) \cdot m_3^{\downarrow K_1 \cap K_3}(C^{\downarrow K_1 \cap K_3})} \cdot \frac{m_2(C^{\downarrow K_2}) \cdot m_3(C^{\downarrow K_3})}{m_3^{\downarrow K_2 \cap K_3}(C^{\downarrow K_2 \cap K_3})} \\
&= \frac{m_1(C^{\downarrow K_1}) \cdot m_2(C^{\downarrow K_2}) \cdot m_3(C^{\downarrow K_3})}{m_2^{\downarrow K_1 \cap K_2}(C^{\downarrow K_1 \cap K_2}) \cdot m_3^{\downarrow K_2 \cap K_3}(C^{\downarrow K_2 \cap K_3})},
\end{aligned}$$

which proves that the equality (1) holds.

C. $C = C^{\downarrow K_1} \otimes C^{\downarrow K_2} \otimes C^{\downarrow K_3}$
& $m_2^{\downarrow K_1 \cap K_2}(C^{\downarrow K_1 \cap K_2}) > 0, m_3^{\downarrow K_2 \cap K_3}(C^{\downarrow K_2 \cap K_3}) = 0.$
In this case, if $C^{\downarrow K_3 \setminus K_2} \neq \mathbf{X}_{K_3 \setminus K_2}$ then both sides of formula (1) equal 0. This is because, due to Definition 1, both composed assignments $(m_1 \triangleright m_2) \triangleright m_3$ and $m_2 \triangleright m_3$ equal 0 for this C, and therefore also $(m_1 \triangleright (m_2 \triangleright m_3))(C) = 0$.

Therefore, consider $C = C^{\downarrow K_1} \otimes C^{\downarrow K_2} \otimes \mathbf{X}_{K_3 \setminus K_2}$. For this we get from Definition 1

$$((m_1 \triangleright m_2) \triangleright m_3)(C) = (m_1 \triangleright m_2)(C^{\downarrow K_1 \cup K_2}).$$

For the right-hand side of formula (1) we get

$$(m_2 \triangleright m_3)(C^{\downarrow K_2 \cup K_3}) = m_2(C^{\downarrow K_2})$$

and therefore

$$(m_1 \triangleright (m_2 \triangleright m_3))(C) = (m_1 \triangleright m_2)(C^{\downarrow K_1 \cup K_2}).$$

D. $C = C^{\downarrow K_1} \otimes C^{\downarrow K_2} \otimes C^{\downarrow K_3}$
& $m_2^{\downarrow K_1 \cap K_2}(C^{\downarrow K_1 \cap K_2}) = 0, m_3^{\downarrow K_2 \cap K_3}(C^{\downarrow K_2 \cap K_3}) > 0.$

focal elements	$(m_1 \triangleright m_2) \triangleright m_3$
$\{a_1 a_2\}$	$\frac{1}{3}$
$\{a_1 \bar{a}_2\}$	$\frac{1}{3}$
$\{a_1 a_2, a_1 \bar{a}_2\}$	$\frac{1}{3}$

Table 1: Composed basic assignment $(m_1 \triangleright m_2) \triangleright m_3$

Since we assume that $m_1^{\downarrow K_1 \cap K_2}(C^{\downarrow K_1 \cap K_2}) > 0$ implies $m_2^{\downarrow K_1 \cap K_2}(C^{\downarrow K_1 \cap K_2}) > 0$, we know that for the considered C, $m_1^{\downarrow K_1 \cap K_2}(C^{\downarrow K_1 \cap K_2}) = 0$, and therefore both sides of formula (1) equal 0 because m_1 is marginal to both $(m_1 \triangleright m_2) \triangleright m_3$ and $m_1 \triangleright (m_2 \triangleright m_3)$.

E. $C = C^{\downarrow K_1} \otimes C^{\downarrow K_2} \otimes C^{\downarrow K_3}$ & $m_2^{\downarrow K_1 \cap K_2}(C^{\downarrow K_1 \cap K_2}) = 0, m_3^{\downarrow K_2 \cap K_3}(C^{\downarrow K_2 \cap K_3}) = 0$. It is obvious from Definition 1 that both sides of formula (1) equal 0 for all C but for $C = C^{\downarrow K_1} \otimes \mathbf{X}_{K_2 \setminus K_1} \otimes \mathbf{X}_{K_3 \setminus K_1}$. For this special case, however,

$$((m_1 \triangleright m_2) \triangleright m_3)(C) = m_1(C^{\downarrow K_1}),$$
$$(m_1 \triangleright (m_2 \triangleright m_3))(C) = m_1(C^{\downarrow K_1}). \qquad \square$$

Example: Let us illustrate the necessity of the assumption

$$m_1^{\downarrow K_1 \cap K_2}(C^{\downarrow K_1 \cap K_2}) > 0 \Longrightarrow m_2^{\downarrow K_1 \cap K_2}(C^{\downarrow K_1 \cap K_2}) > 0$$

required in Lemma 2 by (for the sake of simplicity a rather degenerated) example. Consider three basic assignments m_1, m_2 and m_3. Assume that in this case $K_1 = K_2 = \{1\}$ and $K_3 = \{1, 2\}$, $\mathbf{X}_i = \{a_i, \bar{a}_i\}$ for both $i = 1, 2$. Define $m_1(\{a_1\}) = 1$ and $m_2(\{\bar{a}_1\}) = 1$, which means that both m_1, m_2 have only one focal element, and $m_3(A) = \frac{1}{15}$ for all nonempty subsets of $\mathbf{X}_1 \times \mathbf{X}_2$.

For these basic assignments we immediately get $m_1 = m_1 \triangleright m_2$ (when applying Definition 1, one has to take $C^{\downarrow K_1} \times \mathbf{X}_\emptyset = C^{\downarrow K_1}$), and therefore one gets $m_1 \triangleright m_2 \triangleright m_3$ as indicated in Table 1. Analogously, one gets $m_2 \triangleright m_3$ which is depicted in Table 2. Computing

focal elements	$m_2 \triangleright m_3$
$\{\bar{a}_1 a_2\}$	$\frac{1}{3}$
$\{\bar{a}_1 \bar{a}_2\}$	$\frac{1}{3}$
$\{\bar{a}_1 a_2, a_1 \bar{a}_2\}$	$\frac{1}{3}$

Table 2: Composed basic assignment $m_2 \triangleright m_3$

now the basic assignment $m_1 \triangleright (m_2 \triangleright m_3)$, one gets a basic assignment with only one focal element

$$(m_1 \triangleright (m_2 \triangleright m_3))(\{a_1\} \times \mathbf{X}_2) = 1.$$

Thus we have shown that in this case

$$(m_1 \triangleright m_2) \triangleright m_3 \neq m_1 \triangleright (m_2 \triangleright m_3)\,.$$

3 Decomposable models

3.1 Independence and factorisation

What makes the representation and local computations with multidimensional probability distributions feasible is the property of factorisation [17]. Therefore, in [10] we also introduced this notion into Dempster-Shafer theory of evidence.

Definition 2 ***Simple Factorisation.*** *Consider two nonempty sets $K \cup L = N$. We say that basic assignment m* factorises with respect to (K, L) *if there exist two nonnegative set functions*

$$\phi : \mathcal{P}(\mathbf{X}_K) \longrightarrow [0, +\infty), \quad \psi : \mathcal{P}(\mathbf{X}_L) \longrightarrow [0, +\infty),$$

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

$$m(A) = \begin{cases} \phi(A^{\downarrow K}) \cdot \psi(A^{\downarrow L}) & \textit{if } A = A^{\downarrow K} \otimes A^{\downarrow L} \\ 0 & \textit{otherwise.} \end{cases}$$

Example: Consider $\mathbf{X}_{\{1,2,3\}} = \mathbf{X}_1 \times \mathbf{X}_2 \times \mathbf{X}_3$ with all three $\mathbf{X}_i = \{a_i, \bar{a}_i\}$ as in the preceding example, and consider basic assignment m factorising with respect to $(\{1, 2\}, \{2, 3\})$. This means that it can be represented with the help of two functions

$$\phi : \mathcal{P}(\mathbf{X}_{\{1,2\}}) \to [0, +\infty), \quad \psi : \mathcal{P}(\mathbf{X}_{\{2,3\}}) \to [0, +\infty).$$

Since both subspaces $\mathbf{X}_{\{1,2\}}$ and $\mathbf{X}_{\{2,3\}}$ have 15 nonempty subsets, each of these functions is defined with the help of maximally 15 numbers, which means that the considered basic assignment can be represented with 30 parameters. Generally, a basic assignment on $\mathbf{X}_{\{1,2,3\}}$ can have up to 255 focal elements, and the number of sets $A \subseteq \mathbf{X}_{\{1,2,3\}}$ for which $A \neq A^{\downarrow\{1,2\}} \otimes A^{\downarrow\{2,3\}}$ is 156.

Remark 2 Notice that the importance of the factorisation does not follow only from the fact that the basic assignment m in the preceding example can be represented by two functions ϕ and ψ, i.e., just with 30 parameters, but especially in the fact that the value $m(A)$ can be computed just from two values: $\phi(A^{\downarrow\{1,2\}})$ and $\psi(A^{\downarrow\{2,3\}})$. Value $m(A)$ does not depend on values of functions ϕ and ψ in other points of their domains of definition.

In probability theory, the notion of factorisation is closely connected with the notion of conditional independence. The same holds in Dempster-Shafer theory under the assumption that one accepts the notion of conditional independence as it appears in the following Definition 3, introduced originally in [13]. Nevertheless, based on the recommendation of the anonymous referee, let us first repeat some intuitive reasoning published in [13] that led us to this definition.

There are at least three ways to introduce a generally accepted concept of unconditional (some authors call it marginal) independence (non-interactivity) for two disjoint groups of variables X_K and X_L. Here we will mention two of them, neither of which requires Dempster's rule of combination. The older one, used for example by Ben Yaghlane et al. [1], Shenoy [19] and Studený [21], is based on the properties of a *commonality function* defined for basic assignment m by the formula

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

According to this older definition, we say that disjoint groups of variables X_K and X_L are (unconditionally) independent with respect to basic assignment m if

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

for any $A \subseteq \mathbf{X}_{K \cup L}$. The other (equivalent) definition says that X_K and X_L are independent if for all $A \subseteq \mathbf{X}_{K \cup L}$ for which $A = A^{\downarrow K} \times A^{\downarrow L}$

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

and $m^{\downarrow K \cup L}(A) = 0$ for all the remaining $A \subseteq \mathbf{X}_{K \cup L}$ for which $A \neq A^{\downarrow K} \times A^{\downarrow L}$. Both of these definitions invite generalisation for the case of overlapping groups of variables, both these generalisations satisfy the so-called semigraphoid properties, and yet these generalisations do not coincide. As it is discussed in [2], Studený showed that the generalisation based on the commonality functions is not consistent with marginalisation (for details the reader is referred to [2]), and this is one of the reasons why we prefer the following definition (another reason is that for the concept of conditional independence from Definition 3, one can prove the Factorisation Lemma - see Proposition 3 below).

Definition 3 ***Conditional Independence.*** *Let m be a basic assignment on $\mathbf{X}_N$ and $K, L, M \subset N$ be disjoint, both $K, L \neq \emptyset$. 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 for any $A \subseteq \mathbf{X}_{K \cup L \cup M}$ such that $A = A^{\downarrow K \cup M} \otimes A^{\downarrow L \cup M}$ the equality*

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

holds true, and $m^{\downarrow K \cup L \cup M}(A) = 0$ for all the remaining $A \subseteq \mathbf{X}_{K \cup L \cup M}$, for which $A \neq A^{\downarrow K \cup M} \otimes A^{\downarrow L \cup M}$.

Remark 3 As already mentioned above, it was shown in [13] that this definition meets all the semigraphoid axioms [21] and that for $M = \emptyset$ it reduces to the generally accepted definition of (unconditional, or marginal) independence (see, e.g., [1]).

Important relationships between this type of conditional independence and factorisation (operator of composition) are presented in the following two assertions proved in [14] and [23], respectively.

Proposition 3 ***Factorisation Lemma.*** *Let $K, L \subseteq N$ be nonempty, $K \cup L = N$. m factorises with respect to (K, L) if and only if*

$$K \setminus L \perp\!\!\!\perp L \setminus K \mid K \cap L\ [m].$$

Proposition 4 ***Factorisation of Composition.*** *Let $K, L \subseteq N$ be nonempty, $K \cup L = N$. m factorises with respect to (K, L) if and only if*

$$m = m^{\downarrow K} \triangleright m^{\downarrow L}.$$

3.2 Graphical models

In probability theory, graphical models were defined as probability distributions (measures) factorising with respect to a system of subsets forming cliques of a graph (Daroch, Lauritzen and Speed 1980, Edwards and Havránek 1985). For the sake of this paper we will just define a subclass of graphical models, so-called decomposable models, which factorise with respect to decomposable graphs, i.e., with respect to the graphs whose cliques (maximal complete subsets of nodes) can be ordered to meet the so-called *Running Intersection Property* (RIP): for all $i = 2, \ldots, r$ there exists $j, 1 \leq j < i$, such that

$$K_i \cap (K_1 \cup \ldots \cup K_{i-1}) \subseteq K_j.$$

This offers us a possibility to define decomposable models using Definition 2 recursively.

Definition 4 ***Decomposable Basic Assignments.*** *We say that a basic assignment m is* decomposable *if it* factorises with respect to a decomposable graph *in the following sense (let $K_1, K_2, \ldots, K_r$ be cliques of the considered decomposable graph ordered so that they meet RIP): for all $i = 2, \ldots, r$ the marginal $m^{\downarrow K_1 \cup \ldots \cup K_i}$ factorises (in the sense of Definition 2) with respect to $(K_1 \cup \ldots \cup K_{i-1}, K_i)$.*

By repeated application of Proposition 4 one can see that a decomposable model can easily be represented by a system of its marginals.

Proposition 5 ***Composition of Decomposable Models.*** *Consider a decomposable graph with cliques $K_1, \ldots, K_r$. If this ordering meets RIP then m is decomposable with respect to the graph in question if and only if*

$$m = m^{\downarrow K_1} \triangleright m^{\downarrow K_2} \triangleright \ldots \triangleright m^{\downarrow K_{r-1}} \triangleright m^{\downarrow K_r}.$$

This assertion says that a basic assignment is decomposable if it can be composed from a system of its marginals (the structure of the system must correspond to cliques of a decomposable graph). We can also ask the opposite question: having a system of low-dimensional marginal basic assignment $m_1, m_2, \ldots, m_r$ defined on $\mathbf{X}_{K_1}, \mathbf{X}_{K_2}, \ldots, \mathbf{X}_{K_r}$, respectively, what are the properties of the multidimensional basic assignment $m_1 \triangleright m_2 \triangleright \ldots \triangleright m_r$? The answer to this question, which follows from the following assertion proved in [13], is that if $K_1, K_2, \ldots, K_r$ meet RIP then $m_1 \triangleright m_2 \triangleright \ldots \triangleright m_r$ is decomposable.

Proposition 6 *For any sequence $m_1, m_2, \ldots, m_r$ of basic assignments defined on $\mathbf{X}_{K_1}, \mathbf{X}_{K_2}, \ldots, \mathbf{X}_{K_r}$, respectively, the sequence $\bar{m}_1, \bar{m}_2, \ldots, \bar{m}_r$ computed by the following process*

$$\begin{aligned}
\bar{m}_1 &= m_1, \\
\bar{m}_2 &= \bar{m}_1^{\downarrow K_2 \cap K_1} \triangleright m_2, \\
\bar{m}_3 &= (\bar{m}_1 \triangleright \bar{m}_2)^{\downarrow K_3 \cap (K_1 \cup K_2)} \triangleright m_3, \\
&\vdots \\
\bar{m}_r &= (\bar{m}_1 \triangleright \ldots \triangleright \bar{m}_{r-1})^{\downarrow K_r \cap (K_1 \cup \ldots K_{r-1})} \triangleright m_r,
\end{aligned}$$

has the following properties: $m_1 \triangleright \ldots \triangleright m_r = \bar{m}_1 \triangleright \ldots \triangleright \bar{m}_r$; each $\bar{m}_i$ is defined on $\mathbf{X}_{K_i}$ and is marginal to $m_1 \triangleright \ldots \triangleright m_r$.

Remark 4 It is important to realise that if $K_1, K_2, \ldots, K_r$ meet RIP, then each $K_i \cap (K_1 \cup \ldots K_{i-1})$ is a subset of some K_j $(j < i)$ and therefore

$$(\bar{m}_1 \triangleright \ldots \triangleright \bar{m}_{i-1})^{\downarrow K_i \cap (K_1 \cup \ldots K_{i-1})} = \bar{m}_j^{\downarrow K_i \cap K_j}.$$

Therefore, from the computational point of view, the process described in Proposition 6 is simple for systems of low-dimensional assignments corresponding to decomposable graphs, and can be performed locally (see the next section).

Remark 5 Notice that, thanks to Proposition 3, one can deduce that for a decomposable basic assignment m it is possible to read the system of conditional independence relations valid for m exactly in the same way as it is done for decomposable probabilistic measures: If $G = (N, E)$ is a decomposable graph with respect to which decomposable basic assignment m factorises, and if nodes i and j are separated in G by set M then

$$i \perp\!\!\!\perp j \mid M \; [m].$$

However, let us stress once more: this possibility holds only if one accepts Definition 3.

4 Local computations

By local computations we understand a process based on the ideas published in the famous paper by Lauritzen and Spiegelhalter [17]: the considered probabilistic model (Bayesian network) was first converted into a decomposable model which was subsequently used to compute the required conditional probabilities. What is important in the latter part of the process is the fact that when computing the required conditional probability, one performs computations only on the system of marginal distributions defining the decomposable model. During the computational process one does not need to store more data than what is necessary to store for the decomposable model.

In this section we assume that the considered basic assignment is decomposable, i.e.,

$$m = m^{\downarrow K_1} \triangleright m^{\downarrow K_2} \triangleright \ldots \triangleright m^{\downarrow K_r},$$

and $K_1, K_2, \ldots, K_r$ meet RIP. So let us turn our attention to answering a question: What type of computation will correspond to determination of conditional probability?

Consider the simplest possible case. Assume the goal is to compute a one-dimensional marginal basic assignment for variable X_d in a case where we know that the value of variable X_e equals a $(d, e \in K_1 \cup \ldots \cup K_r)$. If we denote by ${}^{a}_{e}m$ the basic assignment on $\mathbf{X}_e$ with just one focal element ${}^{a}_{e}m(\{a\}) = 1$, then composition ${}^{a}_{e}m \triangleright m$ is a basic assignment describing the situation when one knows that $X_e = a$. Therefore, the goal mentioned above is achieved by computation of $({}^{a}_{e}m \triangleright m)^{\downarrow \{d\}}$.

Now, we are going to study the possibility of computing

$$({}^{a}_{e}m \triangleright m)^{\downarrow\{d\}} = ({}^{a}_{e}m \triangleright (m^{\downarrow K_1} \triangleright m^{\downarrow K_2} \triangleright \ldots \triangleright m^{\downarrow K_r}))^{\downarrow\{d\}}$$

locally. When evaluating $({}^{a}_{e}m \triangleright m)^{\downarrow\{d\}}$ we take full advantage of the assumption that m is decomposable, but, unfortunately, we also have to assume that $\{a\}$ is a focal element of $(m)^{\downarrow\{e\}}$, i.e., $(m)^{\downarrow\{e\}}(\{a\}) > 0$.

Namely, under these assumptions we can make the following consideration:

Having a decomposable model, we can find a permutation of the considered index sets $K_1, K_2, \ldots, K_r$ such that it meets RIP and the sequence starts with any of the sets containing the index e. Without loss of generality, let it be the sequence $K_1, K_2, \ldots, K_r$ (so, $K_1, K_2, \ldots, K_r$ meet RIP and $e \in K_1$). Then we can apply Proposition 2 because $\{e\} \cap K_r \subseteq K_1 \cup \ldots \cup K_{r-1}$ (recall that we selected the ordering such that $e \in K_1$) and

$$(m)^{\downarrow\{e\}}(\{a\}) > 0,$$

from which we get

$$\begin{aligned} &{}^{a}_{e}m \triangleright m \\ &\quad = {}^{a}_{e}m \triangleright ((m^{\downarrow K_1} \triangleright m^{\downarrow K_2} \triangleright \ldots \triangleright m^{\downarrow K_{r-1}}) \triangleright m^{\downarrow K_r}) \\ &\quad = ({}^{a}_{e}m \triangleright (m^{\downarrow K_1} \triangleright m^{\downarrow K_2} \triangleright \ldots \triangleright m^{\downarrow K_{r-1}})) \triangleright m^{\downarrow K_r}. \end{aligned}$$

However, in the same way we also get

$$\begin{aligned} &{}^{a}_{e}m \triangleright (m^{\downarrow K_1} \triangleright m^{\downarrow K_2} \triangleright \ldots \triangleright m^{\downarrow K_{r-1}}) \\ &\quad = ({}^{a}_{e}m \triangleright (m^{\downarrow K_1} \triangleright m^{\downarrow K 2} \triangleright \ldots \triangleright m^{\downarrow K_{r-2}})) \triangleright m^{\downarrow K_{r-1}}, \end{aligned}$$

and after applying Proposition 2 $r-1$ times we get

$${}^{a}_{e}m \triangleright m = {}^{a}_{e}m \triangleright m^{\downarrow K_1} \triangleright m^{\downarrow K_2} \triangleright \ldots \triangleright m^{\downarrow K_{r-1}} \triangleright m^{\downarrow K_r}.$$

So we have shown that if m is a decomposable basic assignment and $(m)^{\downarrow\{e\}}(\{a\}) > 0$, then $({}^{a}_{e}m \triangleright m)^{\downarrow\{d\}}$ can always be computed locally in two steps:

- first order the respective K_i's in the way that they meet RIP and the first K_1 contains index e, and then
- apply Proposition 6 to the decomposable model

$$({}^{a}_{e}m \triangleright m^{\downarrow K_1}) \triangleright m^{\downarrow K_2} \triangleright \ldots \triangleright m^{\downarrow K_r}$$

receiving

$$\begin{aligned} \bar{m}_1 &= {}^{a}_{e}m \triangleright m^{\downarrow K_1}, \\ \bar{m}_2 &= \bar{m}_1^{\downarrow K_2 \cap K_1} \triangleright m^{\downarrow K_2}, \\ \bar{m}_3 &= (\bar{m}_1 \triangleright \bar{m}_2)^{\downarrow K_3 \cap (K_1 \cup K_2)} \triangleright m^{\downarrow K_3}, \\ &\vdots \\ \bar{m}_r &= (\bar{m}_1 \triangleright \ldots \triangleright \bar{m}_{n-1})^{\downarrow K_n \cap (K_1 \cup \ldots K_{n-1})} \triangleright m^{\downarrow K_r}. \end{aligned}$$

focal elements	$m_1(X_1, X_2)$
$\{a_1a_2, a_1\bar{a}_2\}$	$\frac{1}{4}$
$\{a_1\bar{a}_2, \bar{a}_1\bar{a}_2\}$	$\frac{1}{4}$
$\{a_1a_2, a_1\bar{a}_2, \bar{a}_1a_2\}$	$\frac{1}{2}$
	$m_2(X_2, X_3)$
$\{a_2a_3\}$	$\frac{1}{4}$
$\{\bar{a}_2, a_3\}$	$\frac{1}{4}$
$\{a_2\bar{a}_3, \bar{a}_2\bar{a}_3\}$	$\frac{1}{4}$
$\{a_2\bar{a}_3, \bar{a}_2a_3\}$	$\frac{1}{4}$
	$m_3(X_3, X_4)$
$\{a_3a_4\}$	$\frac{1}{2}$
$\{a_3a_4, \bar{a}_3\bar{a}_4\}$	$\frac{1}{4}$
$\{\bar{a}_3a_4, \bar{a}_3\bar{a}_4\}$	$\frac{1}{4}$

Table 3: Basic assignments m_1, m_2, m_3

Now we know that

$${}^{a}_{e}m \triangleright m = \bar{m}_1 \triangleright \bar{m}_2 \triangleright \ldots \triangleright \bar{m}_r,$$

each $\bar{m}_i$ is marginal to ${}^{a}_{e}m \triangleright m$, and therefore the required marginal basic assignment $({}^{a}_{e}m \triangleright m)^{\downarrow\{d\}}$ can be obtained by marginalisation of any m_i for which $d \in K_i$. Recall that, due to RIP, all the computations can be performed locally (see also Remark 4).

Example: Consider a 4-dimensional binary space $\mathbf{X}_1 \times \mathbf{X}_2 \times \mathbf{X}_3 \times \mathbf{X}_4$ with $\mathbf{X}_i = \{a_i, \bar{a}_i\}$, and three two-dimensional basic assignments whose all focal elements are given in Table 3. Let the goal be to compute $(m_1 \triangleright m_2 \triangleright m_3)^{\downarrow\{4\}}$ under the assumption that $X_1 = a_1$, i.e., we want to evaluate

$$({}^{a_1}_{1}m \triangleright (m_1 \triangleright m_2 \triangleright m_3))^{\downarrow\{4\}}.$$

Since X_1 is among the arguments of m_1, and $\{a_1\}$ is a focal element of $(m_1 \triangleright m_2 \triangleright m_3)^{\downarrow\{4\}}$, we can apply the above-introduced procedure (repeated application of Proposition 2) getting that

$$({}^{a_1}_{1}m \triangleright (m_1 \triangleright m_2 \triangleright m_3))^{\downarrow\{4\}} = ({}^{a_1}_{1}m \triangleright m_1 \triangleright m_2 \triangleright m_3)^{\downarrow\{4\}}.$$

So, it remains to apply the process described in Proposition 6. We get that ${}^{a_1}_{1}m \triangleright m_1$ has only one focal element $(\{a_1a_2, a_1\bar{a}_2\})$, and therefore the same holds also for $({}^{a_1}_{1}m \triangleright m_1)^{\downarrow\{2\}}$: $({}^{a_1}_{1}m \triangleright m_1)^{\downarrow\{2\}}(\mathbf{X}_2) = 1$.

From this we immediately get $({}^{a_1}_{1}m \triangleright m_1)^{\downarrow\{2\}} \triangleright m_2$ with two focal elements

$$(({}^{a_1}_{1}m \triangleright m_1)^{\downarrow\{2\}} \triangleright m_2)(\mathbf{X}_2 \times \{\bar{a}_3\}) = \frac{1}{2}$$
$$(({}^{a_1}_{1}m \triangleright m_1)^{\downarrow\{2\}} \triangleright m_2)(\mathbf{X}_2 \times \mathbf{X}_3) = \frac{1}{2},$$

and therefore also its marginal $(({}^{a_1}_{1}m \triangleright m_1)^{\downarrow\{2\}} \triangleright m_2)^{\downarrow\{3\}}$, which is necessary for the computation of the next (already the last) composition, has two focal elements: $\{\bar{a}_3\}$ and $\mathbf{X}_3$. Evaluating this third composition we get that $(({}^{a_1}_{1}m \triangleright m_1)^{\downarrow\{2\}} \triangleright m_2)^{\downarrow\{3\}} \triangleright m_3$ has again two focal elements $\{a_3a_4, \bar{a}_3\bar{a}_4\}$ and $\{\bar{a}_3a_4, \bar{a}_3\bar{a}_4\}$; for each of them the computed composed basic assignment equals $\frac{1}{2}$. Marginalising the last two-dimensional basic assignment we get the desired result:

$$\begin{aligned}&({}^{a_1}_{1}m \triangleright (m_1 \triangleright m_2 \triangleright m_3))^{\downarrow\{4\}}\\ &\quad = ((({}^{a_1}_{1}m \triangleright m_1)^{\downarrow\{2\}} \triangleright m_2)^{\downarrow\{3\}} \triangleright m_3)^{\downarrow\{4\}}\end{aligned}$$

has only one focal element, namely

$$({}^{a_1}_{1}m \triangleright (m_1 \triangleright m_2 \triangleright m_3))^{\downarrow\{4\}})(\bar{a}_4) = 1.$$

Remark 6 If the goal is to compute a basic assignment for variable X_d under the condition that $X_e = a$ and simultaneously $X_f = b$, then one can first compute the decomposable model ${}^{a}_{e}m \triangleright m = \bar{m}_1 \triangleright \bar{m}_2 \triangleright \ldots \triangleright \bar{m}_r$ by the process described above, and afterwards

$${}^{b}_{f}m \triangleright ({}^{a}_{e}m \triangleright m) = {}^{b}_{f}m \triangleright (\bar{m}_1 \triangleright \bar{m}_2 \triangleright \ldots \triangleright \bar{m}_r)$$

in an analogous way finding a new permutation of $K_1, K_2, \ldots, K_r$ meeting RIP such that the first index set contains f. This time, naturally, we have to assume that $m^{\downarrow\{f\}}(\{b\}) > 0$, too.

5 Conclusions

Inspired by Graphical Markov Models in probability theory, we introduced decomposable models in Dempster-Shafer theory of evidence. For this we used two recently introduced concepts: operator of composition and factorisation.

Based on a *factorisation lemma* it is possible to deduce the fact that the introduced decomposable models possess the same conditional independence structure as their probabilistic counterparts; it can be read from the respective graphs following exactly the same rules as in the probabilistic case. This, however, holds only under the assumption that we accept the definition of conditional independence as presented here in Definition 3. Recall that our papers are not the only ones showing evidence in favour of this definition. As it was already presented in [2], Studený showed that the concept of conditional independence based on application of the conjunctive combination rule is not *consistent with marginalisation*. He found two consistent basic assignments for which there does not exist a common extension manifesting the respective conditional independence (for more details and Studený's example see [2]). Let us stress here once more that Definition 3 does not suffer from this insufficiency.

Nevertheless, it was not the main goal of this paper to support the new concept of conditional independence. Here we dealt with the question of whether the ideas of local computations can also be applied to computations in Dempster-Shafer theory of evidence. At this time we have, unfortunately, obtained only a partial answer. The results presented in the last section show that we are able to theoretically support local computations in the cases when the associativity of the operator of composition holds. We did it under the additional assumption that $m^{\downarrow e}(\{a\}) > 0$, i.e., under the assumption that

$$Bel(X_e = a) = m^{\downarrow e}(\{a\}) > 0.$$

From the point of view of real-world application, we would prefer if the designed computational process were applicable under a weaker condition, for example, in a case where

$$Pl(X_e = a) = \sum_{A \subseteq \mathbf{X}_e : a \in A} m^{\downarrow e}(A) > 0.$$

However, as we showed in Example in Section 2, this condition does not guarantee the associativity of the operator of composition. Therefore, there remains an open problem for the further research: either to show that the proposed (or similar) computational process corresponding to local computations can be performed without the assumption of associativity, or to modify the definition of the operator of composition (here we have in mind modification of case [b] of Definition 1) so that associativity would be valid under weaker conditions.

Acknowledgements

This work was supported by GAČR under the grants no. ICC/08/E010, and 201/09/1891, and by MŠMT ČR under grants 1M0572 and 2C06019.

References

[1] B. Ben Yaghlane, Ph. Smets, and K. Mellouli, "Belief Function Independence: I. The Marginal Case," *Int. J. of Approximate Reasoning*, vol. 29, no. 1, pp. 47–70, 2002.

[2] B. Ben Yaghlane, Ph. Smets, and K. Mellouli, "Belief Function Independence: II. The Conditional Case," *Int. J. of Approximate Reasoning*, vol. 31, no. (1-2), pp. 31–75, 2002.

[3] I. Couso, S. Moral and P. Walley, "Examples of independence for imprecise probabilities," in *Proceedings of ISIPTA'99,* G. de Cooman, F. G. Cozman, S. Moral, P. Walley, Eds., Ghent, 1999, pp. 121–130.

[4] J. N. Daroch, S. Lauritzen and T. P. Speed, "Markov Fields and Log Linear Interaction Models for Contingency Tables," *Ann. Stat.* vo. 8, pp. 522-539, 1980.

[5] A. Dempster, "Upper and lower probabilities induced by a multi-valued mapping," *Annals of Math. Statistics* vol. 38, pp. 325–339, 1967.

[6] D. E. Edwards and T. Havránek, "A Fast Procedure for Model Search in Multidimensional Contingency Tables," *Biometrika*, vol. 72, no. 2, pp. 339-351, 1985.

[7] F. V. Jensen, *Bayesian Networks and Decision Graphs.* IEEE Computer Society Press, New York, 2001.

[8] R. Jiroušek, "Composition of probability measures on finite spaces," *Proc. of the 13th Conf. Uncertainty in Artificial Intelligence UAI'97*, (D. Geiger and P. P. Shenoy, eds.). Morgan Kaufmann Publ., San Francisco, California, pp. 274–281, 1997.

[9] R. Jiroušek, "On a conditional irrelevance relation for belief functions based on the operator of composition," in *Dynamics of Knowledge and Belief, Proceedings of the Workshop at the 30th Annual German Conf. on Artificial Intelligence,* Ch. Beierle, G. Kern-Isberner, Eds., Fern Universität in Hagen, Osnabrück, 2007, pp. 28-41.

[10] R. Jiroušek, "Factorization and Decomposable Models in Dempster-Shafer Theory of Evidence," in *Proceedings of the Workshop on Theory of Belief Functions,* Brest, 2010.

[11] R. Jiroušek, "Is It Possible to Define Graphical Models in Dempster-Shafer Theory of Evidence?" in: *Proceedings of the 13th Int. Workshop on Non-Monotonic Reasoning,* Toronto, 2010.

[12] R. Jiroušek, "An Attempt to Define Graphical Models in Dempster-Shafer Theory of Evidence," in: *Proceedings of the 5th International Conference on Soft Methods in Probability and Statistics,* 2010, pp. 361–368.

[13] R. Jiroušek and J. Vejnarová, "Compositional models and conditional independence in evidence theory," *Int. J. Approx. Reasoning*, **52** (2011), 3, pp. 316–334.

[14] R. Jiroušek, J. Vejnarová and M. Daniel, "Compositional models of belief functions," in *Proc. of the 5th Symp. on Imprecise Probabilities and Their Applications,* G. de Cooman, J. Vejnarová, M. Zaffalon, Eds., Praha, 2007, pp. 243–252.

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

[16] S. L. Lauritzen, *Graphical models.* Oxford University Press, 1996.

[17] Lauritzen S. L. and Spiegelhalter D. J., Local computation with probabilities on graphical structures and their application to expert systems, *Journal of the Royal Statistical Society series B* **50** (1988), pp. 157–224.

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

[19] P. P. Shenoy, "Conditional independence in valuation-based systems," *Int. J. of Approximate Reasoning*, vol. 10, no. 3, pp. 203–234, 1994.

[20] P. P. Shenoy, "Binary join trees for computing marginals in the Shenoy-Shafer architecture," *Int. J. of Approximate Reasoning*, vol. 17, no. (2-3), pp. 239–263, 1997.

[21] M. Studený, "Formal properties of conditional independence in different calculi of AI," in *Proceedings of European Conference on Symbolic and quantitative Approaches to Reasoning and Uncertainty ECSQARU'93*, K. Clarke, R. Kruse, S. Moral, Eds., location and date, Springer-Verlag, 1993, pp. 341–351.

[22] M. Studený, "On stochastic conditional independence: the problems of characterization and description," *Annals of Mathematics and Artificial Intelligence,* vol. 35, p. 323-341, 2002.

[23] J. Vejnarová, "On conditional independence in evidence theory," in *Proc. of the 6th Symposium on Imprecise Probability: Theories and Applications,* Durham, UK, 2009, pp. 431–440.