International Journal of Approximate Reasoning 53 (2012) 1155–1167

Contents lists available at SciVerse ScienceDirect

International Journal of Approximate Reasoning

journal homepage: www.elsevier.com/locate/ijar

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

ARTICLE INFO

Article history:
Available online 11 July 2012

Keywords:
Belief network
Composition operator
Conditional independence
Factorisation
Graphical model
Computational complexity

ABSTRACT

When applying any technique of multidimensional models to problems of practice, one always has to cope with two problems: the necessity 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.

1. Introduction

Dempster–Shafer theory of evidence [6,21] generalises classical probability theory in such a way that one can easily describe not only uncertainty but also ignorance. Unfortunately, its disadvantage stems from the fact that belief functions cannot be represented by a point function (like density function in probability theory); instead, one has to manipulate with set functions, which leads to an exponential increase of algorithmic complexity for all the necessary computational procedures.

With regard to probability theory, a 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 the theoretical background was laid by Lauritzen and Spiegelhalter [19]. 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.

By properly studying 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 present a brief survey summarising results concerning decomposable models within Dempster–Shafer theory of evidence presented in [12–14]. 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. The quotation marks in the preceding sentence express the fact that we will consider a very special way of conditioning that can be expressed in the form of a compositional model.

In the rest of this section we introduce necessary notation as well as an operator of composition which plays a crucial role in this paper. Section 2 is devoted to a new property of the operator of composition without which we would not be able to design local computational procedures in Section 5. Section 3 explains the relation between factorisation and the concept of conditional independence (which is different from the one used by most of other authors like Ben Yaglan [3], Shenoy [22]

E-mail address: radim@utia.cas.cz

0888-613X/$ - see front matter
http://dx.doi.org/10.1016/j.ijar.2012.06.012

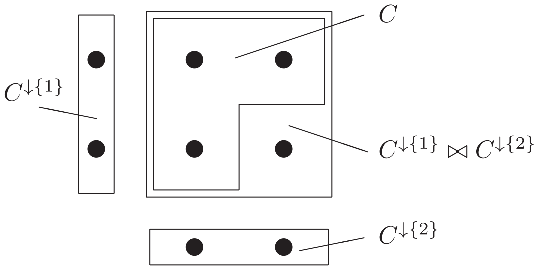

Fig. 1. A set that is not a joint of its projections.

and others), and the above-mentioned survey concerning graphical models is in Section 4. So, most of the assertions from Sections 1, 3, 4 were proved previously and this is why they are presented here without proofs.

1.1. Notation

In this paper we consider a finite multidimensional space $\mathbf{X}_N = \mathbf{X}_1 \times \mathbf{X}_2 \times \cdots \times \mathbf{X}_n$, 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

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

Let us note that if K and L are disjoint, then $A \bowtie B = A \times B$, if $K = L$ then $A \bowtie B = A \cap B$.

From the perspective 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 $C \subseteq C^{\downarrow K} \bowtie C^{\downarrow L}$. However, it does not mean that $C = C^{\downarrow K} \bowtie C^{\downarrow L}$. 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

$$C^{\downarrow\{1\}} \bowtie C^{\downarrow\{2\}} = \{a_1, \bar{a}_1\} \bowtie \{a_2, \bar{a}_2\} = \{(a_1, a_2), (\bar{a}_1, a_2), (a_1, \bar{a}_2), (\bar{a}_1, \bar{a}_2)\} \supsetneq C$$

(see Fig. 1).

1.2. Basic assignments

The role played by a probability distribution in probability theory is in Dempster–Shafer theory played by any of the following set functions: belief function, plausibility function, basic (*probability or belief*) assignment, or commonality function. Knowing one of them, one can derive all the remaining ones. 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

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

and the respective commonality function is

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

Having a basic assignment m on $\mathbf{X}_K$ one can compute 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 [10] as an alternative to Bayesian networks for efficient representation of multidimensional measures. They were based on recurrent application of an operator of composition. This operator is defined for probability measures π and κ on $\mathbf{X}_K$ and $\mathbf{X}_L$, respectively, if the marginal measure $\pi^{\downarrow K\cap L}$ is absolutely continuous with respect to $\kappa^{\downarrow K\cap L}$, for each $x \in \mathbf{X}_{L\cup K}$ by the formula

$$(\pi \triangleright \kappa)(x) = \frac{\pi(x^{\downarrow K})\kappa(x^{\downarrow L})}{\kappa^{\downarrow K\cap L}(x^{\downarrow K\cap L})}$$

(for the precise definition and its properties see [10]). In fact, the operator of composition realises an old Perez' idea [20]: For a probability measure $\pi(x, y, z)$

$$\pi(x, y, z) = \pi(x, y) \cdot \pi(z|x, y)$$

always holds true. It means that if there is not a *strong* conditional dependence between x and z given y, then one can consider probability measure

$$\hat{\pi}(x, y, z) = \pi(x, y) \cdot \pi(z|y)$$

as an approximation of measure π. The advantage of this approximation is that it can easily be reconstructed from two two-dimensional marginals of π. One can immediately see that for measure $\hat{\pi}$, variables x and z are conditionally independent given y. Therefore, Perez called this type of approximation *dependence structure simplification.*

Based on this idea, an analogous operator within the framework of Dempster–Shafer theory was introduced in [17].

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 $C = C^{\downarrow K} \bowtie C^{\downarrow L}$ 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. The reader may have noticed that for C meeting the condition from case [a] (Definition 1), this definition copies the idea of probabilistic composition. Case [b] covers situations when there would appear a positive number divided by zero in the formula from case [a]. In such a situation, the probabilistic operator of composition remains undefined. These are the very situations when basic assignments m_1 and m_2 are in conflict and therefore the whole mass of m_1 is assigned to the respective least informative subset of $\mathbf{X}_{K\cup L}$, i.e., to $C^{\downarrow K} \times \mathbf{X}_{L\setminus K}$. Eventually, case [c] from Definition 1 guarantees that no set $C \neq C^{\downarrow K} \bowtie C^{\downarrow L}$ is assigned a positive mass which would otherwise introduce an undesirable (conditional) dependence.

Remark 2. It is, perhaps, also necessary to stress that the operator of composition is something other than the famous Dempster's rule of combination [6], or its non-normalised version, the so called *conjunctive combination rule* [2]

$$(m_1 \bigcirc\!\!\!\!\cap\, m_2)(C) = \sum_{A \subseteq \mathbf{X}_K, B \subseteq \mathbf{X}_L : A \bowtie B = C} m_1(A) \cdot m_2(B).$$

For example, the operation of composition is (in contrast with the above-mentioned conjunctive combination rule) neither commutative nor associative (see below). 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 itself 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 $(m_1 \circledcap m_2)(C)$ 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 [17]. For its interpretation within the framework of valuation-based systems see [15]. In view of the forthcoming text, the following assertion from [17] is 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 (just consider two basic assignments for which $m_1^{\downarrow K\cap L} \neq m_2^{\downarrow K\cap L}$). 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 $m_1 \triangleright m_2 \triangleright \cdots \triangleright m_r$ defines a (possibly multidimensional) basic assignment defined on $\mathbf{X}_{K_1\cup\cdots\cup K_r}$. Moreover, in contrast to probabilistic case, in D-S theory this composed multidimensional basic assignments is always defined – this is ensured by case [b] of Definition 1.

Example: Consider two basic assignments m_1, m_2 on $\mathbf{X}_{\{1,2\}}, \mathbf{X}_{\{2,3\}}$, respectively, where again each $\mathbf{X}_i = \{a_i, \bar{a}_i\}$. For the sake of simplicity, assume that each of them has only two focal elements, namely: $m_1(\{(a_1, a_2)\}) = 0.5$, $m_1(\{(\bar{a}_1, \bar{a}_2)\}) = 0.5$ and $m_2(\{(a_2, a_3)\}) = 0.6$, $m_2(\{(a_2, \bar{a}_3)\}) = 0.4$. When computing $m_1 \triangleright m_2$, one should realise that although there are 255 nonempty subsets C of $\mathbf{X}_{\{1,2,3\}}$, only 99 of them are such that $C = C^{\downarrow\{1,2\}} \bowtie C^{\downarrow\{2,3\}}$, and Definition 1 assigns positive values only to three of them (case [a] is used twice and case [b] once):

$$\text{[a]}\ \ (m_1 \triangleright m_2)(\{(a_1, a_2, a_3)\}) = \frac{m_1(\{(a_1,a_2)\})\cdot m_2(\{(a_2,a_3)\})}{m_2^{\downarrow\{2\}}(\{(a_2)\})} = \frac{0.5\cdot 0.6}{1} = 0.3,$$

$$\text{[a]}\ \ (m_1 \triangleright m_2)(\{(a_1, a_2, \bar{a}_3)\}) = \frac{m_1(\{(a_1,a_2)\})\cdot m_2(\{(a_2,\bar{a}_3)\})}{m_2^{\downarrow\{2\}}(\{(a_2)\})} = \frac{0.5\cdot 0.4}{1} = 0.2,$$

$$\text{[b]}\ \ (m_1 \triangleright m_2)(\{(\bar{a}_1, \bar{a}_2, a_3), (\bar{a}_1, \bar{a}_2, \bar{a}_3)\}) = m_1(\{(\bar{a}_1, \bar{a}_2)\}) = 0.5.$$

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 $m_1 \triangleright m_2 \triangleright \cdots \triangleright m_r$ means. To avoid the necessity of using too many parentheses, let us make the following convention. In the formulae like $m_1 \triangleright m_2 \triangleright \cdots \triangleright m_r$, 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.,

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

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 [19]), we have to realise why we transfer Bayesian networks into decomposable models. What does make computations in decomposable models easier? The answer is straightforward. Computational procedures have to go through a Bayesian network from source to terminal nodes (parents must be processed before their children). In contrast to this, decomposable models can be reordered so that one can always start with an arbitrary node and the respective computational procedures take this advantage and change the orderings of computations. So it is not surprising that we will need a type of associativity in order to design an efficient computational procedure for compositional models. Surprisingly enough, the following weak form of associativity, which is the main theoretical achievement of this paper, will be sufficient.

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 $K_2 \supseteq K_1 \cap K_3$. If all focal elements of $m_1^{\downarrow K_1\cap K_2}$ are also focal elements of $m_2^{\downarrow K_1\cap K_2}$, i.e.,*

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

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 $C \subseteq \mathbf{X}_{K_1 \cup K_2 \cup K_3}$

$$((m_1 \triangleright m_2) \triangleright m_3)(C) = (m_1 \triangleright (m_2 \triangleright m_3))(C). \tag{1}$$

We have to distinguish five special cases.

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

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

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

B. $C = C^{\downarrow K_1} \bowtie C^{\downarrow K_2} \bowtie C^{\downarrow K_3}$ and $m_2^{\downarrow K_1 \cap K_2}(C^{\downarrow K_1 \cap K_2}) > 0,\ m_3^{\downarrow K_2 \cap K_3}(C^{\downarrow K_2 \cap K_3}) > 0$.
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. $C = C^{\downarrow K_1} \bowtie C^{\downarrow K_2} \bowtie C^{\downarrow K_3}$ and $m_2^{\downarrow K_1 \cap K_2}(C^{\downarrow K_1 \cap K_2}) > 0,\ m_3^{\downarrow K_2 \cap K_3}(C^{\downarrow K_2 \cap K_3}) = 0$.
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 $C = C^{\downarrow K_1} \bowtie C^{\downarrow K_2} \bowtie \mathbf{X}_{K_3 \setminus K_2}$. 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. $C = C^{\downarrow K_1} \bowtie C^{\downarrow K_2} \bowtie C^{\downarrow K_3}$ and $m_2^{\downarrow K_1 \cap K_2}(C^{\downarrow K_1 \cap K_2}) = 0,\ m_3^{\downarrow K_2 \cap K_3}(C^{\downarrow K_2 \cap K_3}) > 0$.
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 $m_1 \triangleright (m_2 \triangleright m_3)$.

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

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 2
Composed basic assignment $m_2 \triangleright m_3$.

Focal elements	$m_2 \triangleright m_3$
$\{(\bar{a}_1, a_2)\}$	$\frac{1}{3}$
$\{(\bar{a}_1, \bar{a}_2)\}$	$\frac{1}{3}$
$\{(\bar{a}_1, a_2), (\bar{a}_1, \bar{a}_2)\}$	$\frac{1}{3}$

E. $C = C^{\downarrow K_1} \bowtie C^{\downarrow K_2} \bowtie C^{\downarrow K_3}$ and $m_2^{\downarrow K_1 \cap K_2}(C^{\downarrow K_1 \cap K_2}) = 0,\ m_3^{\downarrow K_2 \cap K_3}(C^{\downarrow K_2 \cap K_3}) = 0$.
It is obvious from Definition 1 that both sides of formula (1) equal 0 for all C but for $C = C^{\downarrow K_1} \bowtie \mathbf{X}_{K_2 \setminus K_1} \bowtie \mathbf{X}_{K_3 \setminus K_1}$. For this special case, however,
$((m_1 \triangleright m_2) \triangleright m_3)(C) = m_1(C^{\downarrow K_1})$,
$(m_1 \triangleright (m_2 \triangleright m_3))(C) = m_1(C^{\downarrow K_1})$. $\square$

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 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. Independence and factorisation

What makes the representation and local computations with multidimensional probability distributions feasible is the property of factorisation [19], which is closely connected with the notion of (conditional) independence. Already in their seminal papers Dempster [6] and Walley and Fine [27] considered a type of independence that holds for variables X_1 and X_2 with respect to basic assignment m on $\mathbf{X}_{\{1,2\}} = \mathbf{X}_1 \times \mathbf{X}_2$ if for all $A \subseteq \mathbf{X}_{\{1,2\}}$

$$m(A) = \begin{cases} m^{\downarrow\{1\}}(A^{\downarrow\{1\}}) \cdot m^{\downarrow\{2\}}(A^{\downarrow\{2\}}) & \text{if } A = A^{\downarrow\{1\}} \times A^{\downarrow\{2\}}, \\ 0 & \text{otherwise.} \end{cases}$$

This formula inspired us to introduce the following notion of factorisation in Dempster–Shafer theory of evidence [12].

Definition 2 (Simple factorisation)**.** Let m be a basic assignment on $\mathbf{X}_{K \cup L}$ (K, L nonempty). We say that basic assignment m *factorises with respect to the couple* (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 $A \subseteq \mathbf{X}_{K \cup L}$

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

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 examples, and consider basic assignment m factorising with respect to the couple $(\{1,2\},\{2,3\})$. 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 $A \neq A^{\downarrow\{1,2\}} \bowtie A^{\downarrow\{2,3\}}$ is 156.

Remark 3. 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., with 30 parameters, but also from 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. This is important because if we considered basic assignment m on $\mathbf{X}_{\{1,2,3\}}$ that factorises in the sense of the conjunctive combination rule (or Dempster's rule of combination), i.e., there exist basic assignments m_1 and m_2 on $\mathbf{X}_{\{1,2\}}$ and $\mathbf{X}_{\{2,3\}}$, respectively, such that

$$m = m_1 \bigcirc\!\!\!\!\cap\, m_2,$$

then to compute the value $m(A)$ one has to know values of m_1 and m_2 for all supersets of the respective projections of set A.

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 [16]. Nevertheless, let us first repeat some intuitive reasoning published in [16] 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 one used for example by Ben Yaghlane et al. [2], Shenoy [22] and Studený [25] is based on the properties of a *commonality function*. According to this 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, which was already mentioned at the beginning of this section, 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 of them satisfy the so-called semigraphoid properties, both of them are generalisations of the probabilistic notion of conditional independence (i.e., for Bayesian basic assignments they coincide), and yet these generalisations do not coincide in general. As it is discussed in [3], Studený showed that the generalisation based on the commonality functions is not consistent with marginalisation. By this he means that there exist basic assignments m_1 and m_2 on $\mathbf{X}_{\{1,2\}}$ and $\mathbf{X}_{\{2,3\}}$, respectively, for which there exist their common extensions m on $\mathbf{X}_{\{1,2,3\}}$ ($m^{\downarrow\{1,2\}} = m_1$, $m^{\downarrow\{2,3\}} = m_2$), but for none of these extensions X_1 and X_3 are conditionally independent given X_2 (for an example the reader is referred to [3]). 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 $A = A^{\downarrow K\cup M} \bowtie A^{\downarrow L\cup M}$ 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 $A \neq A^{\downarrow K\cup M} \bowtie A^{\downarrow L\cup M}$.

Remark 4. As already mentioned above, it was shown in [11] that this definition meets all the semigraphoid axioms [24] and that for $M = \emptyset$ it reduces to the generally accepted definition of (unconditional) independence (see, e.g., [2]).

Important relationships between this type of conditional independence and factorisation (operator of composition) are presented in the following two assertions proved in [26] and [17], respectively.

Proposition 3 (Factorisation lemma). *Let K, L be nonempty. $m^{\downarrow K\cup L}$ factorises with respect to the couple (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 be nonempty. $m^{\downarrow K\cup L}$ factorises with respect to the couple (K, L) if and only if*

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

Remark 5. It may be interesting to realise that when computing $m^{\downarrow K} \triangleright m^{\downarrow L}$, no positive value $(m^{\downarrow K} \triangleright m^{\downarrow L})(C)$ is assigned by application of the expression [b] of Definition 1. Namely, this expression is applied only when one composes basic assignments which are in conflict, which cannot happen when composing marginals of a more-dimensional basic assignment.

4. Graphical models

4.1. Belief networks

In this section we introduce a Dempster–Shafer counterpart to Bayesian networks. It is well-known that Bayesian networks can be defined in probability theory in several different ways. Here we will proceed according to a rather theoretical approach which defines a Bayesian network as a probability distribution factorising with respect to a given *acyclic directed graph* (DAG). The factorisation guarantees that the independence structure of a probability distribution represented by a Bayesian network is in harmony with the so called *d-separation criterion* [9,18].

The factorisation principle can be formulated in the following way (here $pa(i)$ denotes the set of parents of a node i of the considered DAG, and $fam(i) = pa(i) \cup \{i\}$): measure π is a Bayesian network with a DAG $G = (N, E)$ if for each $i = 2, \ldots, |N|$ (assuming that the ordering $1, 2, \ldots, |N|$ is such that $k \in pa(j) \Longrightarrow k < j$) marginal distribution $\pi^{\downarrow\{1,2,\ldots,i\}}$ factorises with respect to couple $(\{1, 2, \ldots, i-1\}, fam(i))$. And this is the definition which can be directly taken over into Dempster–Shafer theory.

Definition 4 (Belief network). We say that a basic assignment m is a *belief network* (BN) with a DAG $G = (N, E)$ if for each $i = 2, \ldots, |N|$ (assuming the enumeration meets the property that $k \in pa(j) \Longrightarrow k < j$), marginal basic assignment $m^{\downarrow\{1,\ldots,i\}}$ factorises with respect to the couple $(\{1, \ldots, i-1\}, fam(i))$.

From this definition, which differs from those used in [7,23], we get the following description of a BN.

Proposition 5 (Closed form for BN). *Let $G = (N, E)$ be a DAG, and $1, 2, \ldots, |N|$ be its nodes ordered in the way that parents are before their children. Basic assignment m is a BN with graph G if and only if*

$$m = m^{\downarrow fam(1)} \triangleright m^{\downarrow fam(2)} \triangleright \cdots \triangleright m^{\downarrow fam(|N|)}.$$

Proof. Let us employ mathematical induction. For $|N| = 1$ ($fam(1) = \{1\}$) the assertion is trivial, so we will perform the inductive step, which is nothing other than application of Proposition 4: Marginal basic assignment $m^{\downarrow\{1,2,\ldots,i\}}$ factorises with respect to couple $(\{1, 2, \ldots, i-1\}, fam(i))$, and therefore

$$m^{\downarrow\{1,2,\ldots,i\}} = m^{\downarrow\{1,2,\ldots,i-1\}} \triangleright m^{\downarrow fami} = (m^{\downarrow fam(1)} \triangleright \cdots \triangleright m^{\downarrow fam(i-1)}) \triangleright m^{\downarrow fam(i)}. \quad \square$$

Example: With respect to Proposition 5, basic assignment m is a BN with the graph from Fig. 2(a) if

$$m = m^{\downarrow\{1\}} \triangleright m^{\downarrow\{2\}} \triangleright m^{\downarrow\{2,3\}} \triangleright m^{\downarrow\{1,2,4\}} \triangleright m^{\downarrow\{4,5\}} \triangleright m^{\downarrow\{3,5,6\}},$$

because the ordering of nodes 1, 2, 3, 4, 5, 6 is such that parents are before their children. However, it is not the only ordering meeting this condition. Since, say, the ordering 2, 3, 1, 4, 5, 6 also fulfils this condition, the basic assignment m can equivalently be expressed in the form of the following compositional model:

$$m = m^{\downarrow\{2\}} \triangleright m^{\downarrow\{2,3\}} \triangleright m^{\downarrow\{1\}} \triangleright m^{\downarrow\{1,2,4\}} \triangleright m^{\downarrow\{4,5\}} \triangleright m^{\downarrow\{3,5,6\}}.$$

4.2. Factorisation with respect to decomposable graphs

In classical papers on probabilistic models like that by Daroch, Lauritzen and Speed [5], or Edwards and Havránek [8], graphical models were defined as probability distributions (measures) factorising with respect to a system of subsets forming cliques of a graph. 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 $K_1, K_2, \ldots, K_r$ (maximal

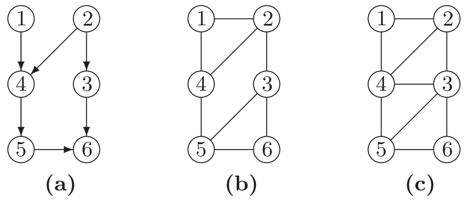

Fig. 2. (a) DAG and (b) decomposable graph.

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, $1 \le j < i$, such that

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

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

Definition 5 (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 $m^{\downarrow K_1 \cup \cdots \cup K_i}$ factorises with respect to the couple $(K_1 \cup \cdots \cup K_{i-1}, K_i)$.

By repeated application of Proposition 4 one can immediately see that a decomposable model can easily be represented by a system of its marginals (the simple proof is in [13]).

Proposition 6 (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*

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

Example: The graph in Fig. 2(c) has four cliques: $\{1, 2, 4\}$, $\{2, 3, 4\}$, $\{3, 4, 5\}$, and $\{3, 5, 6\}$. It is not difficult to verify that this ordering meets the conditions of the running intersection property, which means that the graph is decomposable, and basic assignment m is decomposable with this graph if and only if

$$m = m^{\downarrow\{1,2,4\}} \triangleright m^{\downarrow\{2,3,4\}} \triangleright m^{\downarrow\{3,4,5\}} \triangleright m^{\downarrow\{3,5,6\}},$$

or, using another RIP ordering,

$$m = m^{\downarrow\{2,3,4\}} \triangleright m^{\downarrow\{3,4,5\}} \triangleright m^{\downarrow\{1,2,4\}} \triangleright m^{\downarrow\{3,5,6\}},$$

or,

$$m = m^{\downarrow\{3,5,6\}} \triangleright m^{\downarrow\{3,4,5\}} \triangleright m^{\downarrow\{2,3,4\}} \triangleright m^{\downarrow\{1,2,4\}}.$$

Let us stress that it can be shown that all three of these conditions are equivalent because all three clique orderings considered here do meet RIP. Notice the characteristic property expressed by RIP: whenever an operator of composition is realised the composition is computed, in fact, for two three-dimensional marginals.

Proposition 6 says that a basic assignment is decomposable if and only 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 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, what are the properties of the multidimensional basic assignment $m_1 \triangleright m_2 \triangleright \cdots \triangleright m_r$? The answer to this question, which follows from the following assertion proved in [16], is that if $K_1, K_2, \ldots, K_r$ meet RIP then $m_1 \triangleright m_2 \triangleright \cdots \triangleright m_r$ is decomposable.

Proposition 7. *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*

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

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

Remark 6. It is important to realise that if $K_1, K_2, \ldots, K_r$ meet RIP, then each $K_i \cap (K_1 \cup \cdots \cup K_{i-1})$ is a subset of some K_j $(j < i)$ and therefore

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

Therefore, from the computational point of view, the process described in Proposition 7 is simple for systems of low-dimensional assignments corresponding to decomposable graphs, and can be performed locally (see the next section).

Remark 7. 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.

5. Local computations

By local computations we understand a process based on the ideas published in the famous paper by Lauritzen and Spiegelhalter [19]: The considered probabilistic model (Bayesian network) is first converted into a decomposable model which is 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 paper we do not have an ambition to solve this problem in full generality. We just discuss a way that will enable us to answer a question like: *What is a belief for values of variable X_j if we know that variable X_i has a value a?* As said above, in probability theory the answer is given by conditional probability distribution $\pi(X_j|X_i = a)$. Let us study a possibility to obtain this conditional probability distribution with the help of the probabilistic operator of composition (see the beginning of Section 1.3).

Define a *degenerated* one-dimensional probability distribution $\kappa_{|i;a}$ as a distribution of variable X_i achieving probability 1 for value $X_i = a$, i.e.,

$$\kappa_{|i;a}(X_i = x) = \begin{cases} 1 & \text{if } x = a, \\ 0 & \text{otherwise.} \end{cases}$$

Now, compute $(\kappa_{|i;a} \triangleright \pi)^{\downarrow\{j\}}$ for a probability distribution π of variables X_K with $i, j \in K$:

$$\begin{aligned}(\kappa_{|i;a} \triangleright \pi)^{\downarrow\{j\}}(y) &= ((\kappa_{|i;a} \triangleright \pi)^{\downarrow\{j,i\}})^{\downarrow\{j\}}(y) = (\kappa_{|i;a} \triangleright \pi^{\downarrow\{j,i\}})^{\downarrow\{j\}}(y) \\ &= \sum_{x \in \mathbf{X}_i} \frac{\kappa_{|i;a}(x) \cdot \pi^{\downarrow\{j,i\}}(y, x)}{\pi^{\downarrow\{i\}}(x)} = \frac{\pi^{\downarrow\{j,i\}}(y, a)}{\pi^{\downarrow\{i\}}(a)} = \pi^{\downarrow\{j,i\}}(y|a).\end{aligned}$$

This fact that the conditional probability $\pi^{\downarrow\{j,i\}}(y|a)$ can be expressed with the help of the operator of composition, inspired us to also introduce a similar construction for basic assignments. Define a degenerated basic assignment $m_{|i;a}$ on $\mathbf{X}_i$ with only one focal element $m_{|i;a}(\{a\}) = 1$. What is the basic assignment $(m_{|i;a} \triangleright m)^{\downarrow\{j\}}$? The answer is given by Proposition 1: it is that basic assignment which arises from m by changing its marginal for variable X_i so that it is equal to $m_{|i;a}$. In other words, it describes the relationships among all variables from X_N which are encoded in m when we know that X_i takes the value a. Therefore, in a sense it yields an answer to the question raised above.

In the rest of this section we will show that having a belief network m, it is possible to compute $m_{|i;a} \triangleright m$ by a computational process following the ideas of Lauritzen and Spiegelhalter.

5.1. Conversion of a BN into decomposable basic assignment

The process realizing this step can be directly taken over from probability theory [9]. We start assuming that the considered basic assignment m is given in a form of a belief network, i.e.,

$$m = m^{\downarrow fam(1)} \triangleright m^{\downarrow fam(2)} \triangleright \cdots \triangleright m^{\downarrow fam(|N|)},$$

for an acyclic graph $G = (N, E)$, and the ordering $1, 2, \ldots, |N|$ is such that parents are before their children. Then undirected graph $G = (N, \bar{E})$, where

$$\bar{E} = \left\{ \{i,j\} \in \binom{N}{2} : \exists k \in N \quad \{i,j\} \subseteq fam(k) \right\},$$

is a so-called *moral graph* from which one can get the necessary decomposable graph $G = (V, F)$ (which will be uniquely specified by a system of its cliques $C_1, C_2, \ldots, C_r$) by any heuristic approach used for moral graph triangulation [4] (it is known that the process of looking for an optimal triangulated graph is a NP hard problem). When one realises that there must exist an ordering (let it be the ordering $C_1, C_2, \ldots, C_r$) of the cliques meeting RIP and simultaneously

$$i \in pa(j) \implies f(i) \le f(j),$$

where $f(k) = \min(\ell : k \in C_\ell)$, then it is an easy task to compute the necessary marginal ba's $m^{\downarrow C_1}, \ldots, m^{\downarrow C_r}$.

Example: Consider a basic assignment m that is a BN with the graph in Fig. 2(a). It means that

$$m = m^{\downarrow\{1\}} \triangleright m^{\downarrow\{2\}} \triangleright m^{\downarrow\{2,3\}} \triangleright m^{\downarrow\{1,2,4\}} \triangleright m^{\downarrow\{4,5\}} \triangleright m^{\downarrow\{3,5,6\}},$$

or equivalently it means that the basic assignment m can be represented with the help of two one-dimensional (m_1, m_2), two two-dimensional (m_{23}, m_{45}), and two three-dimensional (m_{124}, m_{356}) basic assignments

$$m = m_1 \triangleright m_2 \triangleright m_{23} \triangleright m_{124} \triangleright m_{45} \triangleright m_{356}.$$

Notice that here we do not assume that, say, m_{124} is a marginal of m.

The corresponding moral graph is in Fig. 2(b), and a possible triangulated (decomposable) graph is in Fig. 2(c). So, the corresponding decomposable model is represented with the help of four three-dimensional marginals, which can be computed in the following way:

$$\begin{aligned} m^{\downarrow\{1,2,4\}} &= m_1 \triangleright m_2 \triangleright m_{124}, \\ m^{\downarrow\{2,3,4\}} &= m^{\downarrow\{2,4\}} \triangleright m_{23}, \\ m^{\downarrow\{3,4,5\}} &= m^{\downarrow\{3,4\}} \triangleright m_{45}, \\ m^{\downarrow\{3,5,6\}} &= m^{\downarrow\{3,5\}} \triangleright m_{356}. \end{aligned}$$

5.2. *Computation of conditional basic assignment*

In comparison with the previous step, this computational process is much more complex. We have to show that having a decomposable basic assignment $m = m^{\downarrow C_1} \triangleright \cdots \triangleright m^{\downarrow C_r}$ one can compute $(m_{|i;a} \triangleright m)^{\downarrow\{j\}}$ locally.

For this, we take advantage of the famous fact (an immediate consequence of the existence of a join tree, see [1]) that if $C_1, C_2, \ldots, C_r$ can be ordered to meet RIP, then for each $k \in \{1, 2, \ldots, r\}$ there exists an ordering meeting RIP for which C_k is the first one. So consider any C_k for which $i \in C_k$, and find the ordering meeting RIP which starts with C_k. Without loss of generality let it be $C_1, C_2, \ldots, C_r$ (so, $i \in C_1$).

Considering basic assignment m decomposable with respect to the graph with cliques $C_1, C_2, \ldots, C_r$, our goal is to compute

$$(m_{|i;a} \triangleright m)^{\downarrow\{j\}} = \left(m_{|i;a} \triangleright (m^{\downarrow C_1} \triangleright m^{\downarrow C_2} \triangleright \cdots \triangleright m^{\downarrow C_r})\right)^{\downarrow\{j\}}.$$

However, at this moment we have to assume that $m^{\downarrow\{i\}}(\{a\})$ is positive. Under this assumption we can apply Proposition 2 $(r-1)$ times getting

$$\begin{aligned} m_{|i;a} \triangleright (m^{\downarrow C_1} \triangleright m^{\downarrow C_2} \triangleright \cdots \triangleright m^{\downarrow C_r}) &= m_{|i;a} \triangleright (m^{\downarrow C_1} \triangleright m^{\downarrow C_2} \triangleright \cdots \triangleright m^{\downarrow C_{r-1}}) \triangleright m^{\downarrow C_r} \\ &= \cdots = m_{|i;a} \triangleright m^{\downarrow C_1} \triangleright m^{\downarrow C_2} \triangleright \cdots \triangleright m^{\downarrow C_r}, \end{aligned}$$

from which the following computationally local process (see Remark 6)

$$\begin{aligned} \bar{m}_1 &= m_{|i;a} \triangleright m^{\downarrow C_1}, \\ \bar{m}_2 &= \bar{m}_1^{\downarrow C_2 \cap C_1} \triangleright m^{\downarrow C_2}, \\ \bar{m}_3 &= (\bar{m}_1 \triangleright \bar{m}_2)^{\downarrow C_3 \cap (C_1 \cup C_2)} \triangleright m^{\downarrow C_3}, \\ &\vdots \\ \bar{m}_r &= (\bar{m}_1 \triangleright \cdots \triangleright \bar{m}_{r-1})^{\downarrow C_r \cap (C_1 \cup \cdots C_{r-1})} \triangleright m^{\downarrow C_r}, \end{aligned}$$

Table 3
Focal elements of basic assignments m_1, m_2, m_3.

$m_1(\{(a_1, a_2), (a_1, \bar{a}_2)\}) = \frac{1}{4}$	$m_2(\{(a_2, a_3)\}) = \frac{1}{4}$	$m_3(\{(a_3, a_4)\}) = \frac{1}{2}$
$m_1(\{(a_1, \bar{a}_2), (\bar{a}_1, \bar{a}_2)\}) = \frac{1}{4}$	$m_2(\{(\bar{a}_2, a_3)\}) = \frac{1}{4}$	$m_3(\{(a_3, a_4), (\bar{a}_3, \bar{a}_4)\}) = \frac{1}{4}$
$m_1(\{(a_1, a_2), (a_1, \bar{a}_2), (\bar{a}_1, a_2)\}) = \frac{1}{2}$	$m_2(\{(a_2, \bar{a}_3), (\bar{a}_2, \bar{a}_3)\}) = \frac{1}{4}$	$m_3(\{(\bar{a}_3, a_4), (\bar{a}_3, \bar{a}_4)\}) = \frac{1}{4}$
	$m_2(\{(a_2, \bar{a}_3), (\bar{a}_2, a_3)\}) = \frac{1}{4}$	

yields a sequence $\bar{m}_1, \ldots, \bar{m}_r$, such that $m_{|i;a} \triangleright m = \bar{m}_1 \triangleright \cdots \triangleright \bar{m}_r$, and each $\bar{m}_k = (m_{|i;a} \triangleright m)^{\downarrow C_k}$. Therefore, to compute $(m_{|i;a} \triangleright m)^{\downarrow \{j\}}$ it is enough to find any k such that $j \in C_k$ because in this case $(m_{|i;a} \triangleright m)^{\downarrow \{j\}} = \bar{m}_k^{\downarrow \{j\}}$.

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

$$(m_{|1;a_1} \triangleright (m_1 \triangleright m_2 \triangleright m_3))^{\downarrow \{4\}}.$$

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

$$(m_{|1;a_1} \triangleright (m_1 \triangleright m_2 \triangleright m_3))^{\downarrow \{4\}} = (m_{|1;a_1} \triangleright m_1 \triangleright m_2 \triangleright m_3)^{\downarrow \{4\}}.$$

So, the task remains to apply the process described in Proposition 7. We get that $m_{|1;a_1} \triangleright m_1$ has only one focal element $(\{(a_1, a_2), (a_1, \bar{a}_2)\})$, and therefore the same holds also for $(m_{|1;a_1} \triangleright m_1)^{\downarrow \{2\}}$: $(m_{|1;a_1} \triangleright m_1)^{\downarrow \{2\}}(\mathbf{X}_2) = 1$.

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

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

and therefore also its marginal $((m_{|1;a_1} \triangleright m_1)^{\downarrow \{2\}} \triangleright m_2)^{\downarrow \{3\}}$, 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 $((m_{|1;a_1} \triangleright m_1)^{\downarrow \{2\}} \triangleright m_2)^{\downarrow \{3\}} \triangleright m_3$ has again two focal elements $\{(a_3, a_4), (\bar{a}_3, \bar{a}_4)\}$ and $\{(\bar{a}_3, a_4), (\bar{a}_3, \bar{a}_4)\}$; 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:

$$(m_{|1;a_1} \triangleright (m_1 \triangleright m_2 \triangleright m_3))^{\downarrow \{4\}} = (((m_{|1;a_1} \triangleright m_1)^{\downarrow \{2\}} \triangleright m_2)^{\downarrow \{3\}} \triangleright m_3)^{\downarrow \{4\}}$$

has only one focal element, namely

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

Remark 8. 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 $m_{|e;a} \triangleright m = \bar{m}_1 \triangleright \bar{m}_2 \triangleright \cdots \triangleright \bar{m}_r$ by the process described above, and afterwards

$$m_{|f;b} \triangleright (m_{|e;a} \triangleright m) = m_{|f;b} \triangleright (\bar{m}_1 \triangleright \bar{m}_2 \triangleright \cdots \triangleright \bar{m}_r)$$

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.

6. 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 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 [3], 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 [3]). 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 necessary associativity of the operator of composition.

In this paper we studied the possibility to compute a posterior basic assignment under a condition that a value of a variable is given. But it should be mentioned that the described procedure is applicable also in case one wants to compute conditional basic assignment like, e.g., that studied by Shenoy in [22]. In fact it can be used for computation of any basic assignment that can be expressed as a composition of a specific (perhaps one-dimensional) assignment with a multidimensional decomposable model.

Acknowledgements

This work was supported by GAČR under the grant no. 403/12/2175.

References

[1] C. Beeri, R. Fagin, D. Maier, M. Yannakakis, On the desirability of acyclic database schemes, J. ACM 30 (3) (1983) 479–513.
[2] B. Ben Yaghlane, Ph. Smets, K. Mellouli, Belief function independence: I. The marginal case, Int. J. Approx. Reason. 29 (1) (2002) 47–70.
[3] B. Ben Yaghlane, Ph. Smets, K. Mellouli, Belief function independence: II. The conditional case, Int. J. Approx. Reason. 31 (1–2) (2002) 31–75.
[4] A. Cano, S. Moral, Heuristic algorithms for the triangulation of graphs, in: B. Bouchon-Meunier, R.R. Yager, L.A. Zadeh (Eds.), Advances in Intelligent Computing-IPMU '94, Paris, France, 1995.
[5] J.N. Daroch, S. Lauritzen, T.P. Speed, Markov fields and log linear interaction models for contingency tables, Ann. Stat. 8 (1980) 522–539.
[6] A.P. Dempster, Upper and lower probabilities induced by a multi-valued mapping, Ann. Math. Stat. 38 (1967) 325–339.
[7] A.P. Dempster, A. Kong, Uncertain evidence and artificial analysis, J. Stat. Plann. Infer. 20 (1988) 355–368.
[8] D.E. Edwards, T. Havránek, A fast procedure for model search in multidimensional contingency tables, Biometrika 72 (2) (1985) 339–351.
[9] F.V. Jensen, Bayesian Networks and Decision Graphs, IEEE Computer Society Press, New York, 2001.
[10] R. Jiroušek, Composition of probability measures on finite spaces, in: D. Geiger, P.P. Shenoy (Eds.), Proceedings of the 13th Conference Uncertainty in Artificial Intelligence UAI'97, Morgan Kaufman Publ., San Francisco, California, 1997, pp. 274–281.
[11] R. Jiroušek, On a conditional irrelevance relation for belief functions based on the operator of composition, in: Ch. Beierle, G. Kern-Isberner (Eds.), Dynamics of Knowledge and Belief, Proceedings of the Workshop at the 30th Annual German Conf. on Artificial Intelligence, Fern Universität in Hagen, Osnabrück, 2007, pp. 28–41.
[12] R. Jiroušek, Factorization and decomposable models in Dempster–Shafer theory of evidence, in: Proceedings of the Workshop on Theory of Belief Functions, Brest, 2010.
[13] R. Jiroušek, Is it possible to define graphical models in Dempster–Shafer theory of evidence? in: Proceedings of the 13th International Workshop on Non-Monotonic Reasoning, Toronto, 2010.
[14] 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.
[15] R. Jiroušek, P.P. Shenoy, Compositional models in valuation-based systems, Working Paper No. 325, School of Business, University of Kansas, Lawrence, KS, 2011.
[16] R. Jiroušek, J. Vejnarová, Compositional models and conditional independence in evidence theory, Int. J. Approx. Reason. 52 (3) (2011) 316–334.
[17] R. Jiroušek, J. Vejnarová, M. Daniel, Compositional models of belief functions, in: G. de Cooman, J. Vejnarová, M. Zaffalon (Eds.), Proceedings of the Fifth International Symposium on Imprecise Probability: Theories and Applications, Praha, 2007, pp. 243–252.
[18] S.L. Lauritzen, Graphical Models, Oxford University Press, 1996.
[19] S.L. Lauritzen, D.J. Spiegelhalter, Local computation with probabilities on graphical structures and their application to expert systems, J. Roy. Stat. Soc. Ser. B 50 (1988) 157–224.
[20] A. Perez, ε-Admissible simplification of the dependence structure of a set of random variables, Kybernetika 13 (1977) 439–449.
[21] G. Shafer, A Mathematical Theory of Evidence, Princeton University Press, Princeton, NJ, 1976.
[22] P.P. Shenoy, Conditional independence in valuation-based systems, Int. J. Approx. Reason. 10 (3) (1994) 203–234.
[23] P.P. Shenoy, G. Shafer, Axioms for probability and belief-function propagation, in: R.D. Shachter, T. Levitt, J.F. Lemmer, L.N. Kanal (Eds.), Uncertainty in Artificial Intelligence, vol. 4, North-Holland, 1990.
[24] M. Studený, Formal properties of conditional independence in different calculi of AI, in: K. Clarke, R. Kruse, S. Moral (Eds.), Proceedings of European Conference on Symbolic and quantitative Approaches to Reasoning and Uncertainty ECSQARU'93, Springer-Verlag, 1993, pp. 341–351.
[25] M. Studený, On stochastic conditional independence: the problems of characterization and description, Ann. Math. Artif. Intell. 35 (2002) 323–341.
[26] J. Vejnarová, On conditional independence in evidence theory, in: Proceedings of the 6th Symposium on Imprecise Probability: Theories and Applications Durham, UK, 2009, pp. 431–440.
[27] P. Walley, T.L. Fine, Towards a frequentist theory of upper and lower probability, Ann. Stat. 10 (1982) 749–761.