Compositional Belief Function Models

Radim Jiroušek
Faculty of Management, Jindřichův Hradec
University of Economics, Prague
and
Institute of Information Theory and Automation of the ASCR
Czech Republic
Email: radim@utia.cas.cz

***Abstract*—Analogously to Graphical Markov models, also Compositional models serve as an efficient tool for multidimensional models representation. The main idea of the latter models resembles a jig saw puzzle: Multidimensional models are assembled (composed) from a large number of small pieces, from a large number of low-dimensional models. Originally they were designed to represent multidimensional probability distributions. In this paper they will be used to represent multidimensional belief functions (or more precisely, multidimensional basic belief assignments) with the help of a system of low-dimensional ones.**

In addition to a number of basic properties of such models, in the paper it will be shown that these models can serve as a real enrichment of probabilistic models. They can relieve a drawback of probabilistic models that can be, in case that the initial building blocks of the model are inconsistent, undefined. As a side result of the paper we propose a new way how to define the concept of conditional independence for belief functions.

I. INTRODUCTION

The main bottleneck of applications of belief function models to problems of practice lies in the fact that a belief measure, in contrast to a probability or possibility measure, cannot be represented by a density function. It is a set function and for its representation one needs an exponential number of parameters (exponential with the size of a finite space on which the belief function measure is defined). Therefore, one has to employ some approaches enabling reduction of necessary parameters. In this contribution we will discuss an approach utilizing properties of conditional independence relations that will enable us to assemble (*compose*) a multidimensional model from a system of its marginal submodels. This is also the reason why these models are called *compositional models*.

II. SET PROJECTIONS AND EXTENSIONS

In the whole paper we shall deal with a finite number of variables $X_1, X_2, \ldots, X_n$ each of which is specified by a finite set $\mathbf{X}_i$ of its values. So, we will consider multidimensional space (in the belief function setting it is usually called *frame of discernment*)

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

and its *subspaces*. For $K \subset N = \{1, 2, \ldots, n\}$, $\mathbf{X}_K$ denotes a Cartesian product of those $\mathbf{X}_i$, for which $i \in K$:

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

and $X_K = \{X_i\}_{i \in K}$ denotes the set of the respective variables.

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

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

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

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

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

In addition to the projection, in this text we will need also an opposite operation which will be called extension. By the *extension* of two sets $A \subseteq \mathbf{X}_{K_1}$ and $B \subseteq \mathbf{X}_{K_2}$ we will understand a set

$$A \otimes B = \{x \in \mathbf{X}_{K_1 \cup K_2} : x^{\downarrow K_1} \in A \ \ \& \ \ x^{\downarrow K_2} \in B\}.$$

Notice that if K_1 and K_2 are disjoint then their extension is just their Cartesian product

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

If $K_1 \cap K_2 \neq \emptyset$ and $A^{\downarrow K_1 \cap K_2} \cap B^{\downarrow K_1 \cap K_2} = \emptyset$ then also $A \otimes B = \emptyset$.

In what follows, an important role will be played by special sets, which were in [2] called Z-layered rectangles. These are those sets $C \subseteq \mathbf{X}_{K_1 \cup K_2}$ for which

$$C = C^{\downarrow K_1} \otimes C^{\downarrow K_2}.$$

III. COMPOSITION OF BASIC ASSIGNMENTS

A belief function is defined with the help of a *basic* (*probability or belief*) *assignment* m on $\mathbf{X}_N$, which is a set function

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

with

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

Therefore, for the sake of simplicity, we will not speak about belief functions but about basic assignments: We shall marginalize and compose basic assignments. For each $K \subset N$

marginal basic assignment of m is defined (for each $B \subseteq \mathbf{X}_K$):

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

Notice that we again do not exclude situation when K is empty. In this case we get $m^{\downarrow \emptyset}(\emptyset) = 1$. This is important with respect to the following notion of projectiveness: We say that two basic assignments m_1 and m_2 defined on $\mathbf{X}_{K_1}$ and $\mathbf{X}_{K_2}$, respectively, are *projective* if

$$m_1^{\downarrow K_1 \cap K_2} = m_2^{\downarrow K_1 \cap K_2}.$$

Now, we can define the most important notion of this paper, which was originally defined in [5].

Definition 1: For arbitrary two basic assignments m_1 on $\mathbf{X}_{K_1}$ and m_2 on $\mathbf{X}_{K_2}$ ($K_1 \neq \emptyset \neq K_2$) a *composition* $m_1 \triangleright m_2$ is defined for each $C \subseteq \mathbf{X}_{K_1 \cup K_2}$ by one of the following expressions:

[a] if $m_2^{\downarrow K_1 \cap K_2}(C^{\downarrow K_1 \cap K_2}) > 0$ and $C = C^{\downarrow K_1} \otimes C^{\downarrow K_2}$ then

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

[b] if $m_2^{\downarrow K_1 \cap K_2}(C^{\downarrow K_1 \cap K_2}) = 0$ and $C = C^{\downarrow K_1} \times \mathbf{X}_{K_2 \setminus K_1}$ then

$$(m_1 \triangleright m_2)(C) = m_1(C^{\downarrow K_1});$$

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

Before we start studying formal properties of this operator let us illustrate both marginalization and composition on a simple example.

Example 1: Consider three binary variables X_1, X_2, X_3 with $\mathbf{X}_1 = \{a, \bar{a}\}$, $\mathbf{X}_2 = \{b, \bar{b}\}$, $\mathbf{X}_3 = \{c, \bar{c}\}$, and two 2-dimensional basic assignments m_1 and m_2 as specified in Table I.

Notice that these two assignments are not projective; for this see their one-dimensional marginals in Table II. Therefore, because of property (3) of Lemma 1 presented below, $m_1 \triangleright m_2 \neq m_2 \triangleright m_1$.

How difficult is to compute such a composition? To determine general 3-dimensional assignment (of binary variables) one has to specify 255 numbers, because $\mathbf{X}_{\{1,2,3\}}$ has $2^8 - 1 = 255$ nonempty subsets. However, when computing $m_1 \triangleright m_2$, most of these 255 values equal 0 because most of these subsets do not meet the condition $C = C^{\downarrow\{1,2\}} \otimes C^{\downarrow\{2,3\}}$ and therefore the corresponding value of the assignment $m_1 \triangleright m_2$ is defined by the point [c] of the definition.

What are the subsets for which $C \neq C^{\downarrow\{1,2\}} \otimes C^{\downarrow\{2,3\}}$? For example, it is easy to show that all the sets of cardinality 7 belong to this category (hint: show that for any $C \subseteq \mathbf{X}_{\{1,2,3\}}$, for which $|C| = 7$, $C^{\downarrow\{1,2\}} = \mathbf{X}_{\{1,2\}}$ and $C^{\downarrow\{2,3\}} = \mathbf{X}_{\{2,3\}}$).

Since all singletons (one-point-sets) meet the considered equality $C = C^{\downarrow\{1,2\}} \otimes C^{\downarrow\{2,3\}}$, all sets C, for which $C \neq C^{\downarrow\{1,2\}} \otimes C^{\downarrow\{2,3\}}$ must have at least two elements: an example is $\{abc, \bar{a}b\bar{c}\}$. As further examples may serve sets $\{ab\bar{c}, \bar{a}bc, \bar{a}b\bar{c}, a\bar{b}c\}$ and $\{\bar{a}bc, a\bar{b}c, ab\bar{c}\}$. A common characteristics of all these sets is that assigning a positive belief to them one introduces a type of conditional relationship between X_1 and X_3 given (at least one) value of X_2.

Let us turn our attention back to the computation of $m_1 \triangleright m_2$ for the assignments of our example. Since $m_2^{\downarrow\{2\}}(b) = 0$, one immediately notices that point [b] of the definition is used whenever $C \subseteq \mathbf{X}_{\{1,2,3\}}$ is considered for which $C^{\downarrow\{2\}} = b$. Considering all what has been said we get only 8 subsets, for which assignment $m_1 \triangleright m_2$ is positive - see Table III, where the first column bears the information, which point of the definition is used to compute the respective value.

TABLE I
BASIC ASSIGNMENTS $m_1(x_{\{1,2\}})$ AND $m_2(x_{\{2,3\}})$.

$C \subseteq \mathbf{X}_{\{1,2\}}$	$m_1(C)$
$\{ab\}$	0.1
$\{a\bar{b}\}$	0.5
$\{\bar{a}b\}$	0.2
$\{\bar{a}\bar{b}\}$	0
$\{ab, a\bar{b}\}$	0
$\{ab, \bar{a}b\}$	0
$\{ab, \bar{a}\bar{b}\}$	0
$\{a\bar{b}, \bar{a}b\}$	0
$\{a\bar{b}, \bar{a}\bar{b}\}$	0
$\{\bar{a}b, \bar{a}\bar{b}\}$	0
$\{ab, a\bar{b}, \bar{a}b\}$	0
$\{ab, a\bar{b}, \bar{a}\bar{b}\}$	0
$\{ab, \bar{a}b, \bar{a}\bar{b}\}$	0
$\{a\bar{b}, \bar{a}b, \bar{a}\bar{b}\}$	0
$\{ab, a\bar{b}, \bar{a}b, \bar{a}\bar{b}\}$	0.2

$C \subseteq \mathbf{X}_{\{2,3\}}$	$m_2(C)$
$\{bc\}$	0
$\{b\bar{c}\}$	0
$\{\bar{b}c\}$	0.3
$\{\bar{b}\bar{c}\}$	0.1
$\{bc, b\bar{c}\}$	0
$\{bc, \bar{b}c\}$	0
$\{bc, \bar{b}\bar{c}\}$	0.1
$\{b\bar{c}, \bar{b}c\}$	0
$\{b\bar{c}, \bar{b}\bar{c}\}$	0
$\{\bar{b}c, \bar{b}\bar{c}\}$	0.1
$\{bc, b\bar{c}, \bar{b}c\}$	0
$\{bc, b\bar{c}, \bar{b}\bar{c}\}$	0
$\{bc, \bar{b}c, \bar{b}\bar{c}\}$	0.3
$\{b\bar{c}, \bar{b}c, \bar{b}\bar{c}\}$	0
$\{bc, b\bar{c}, \bar{b}c, \bar{b}\bar{c}\}$	0.1

TABLE II
ONE-DIMENSIONAL MARGINAL ASSIGNMENTS $m_1^{\downarrow\{2\}}$ AND $m_2^{\downarrow\{2\}}$.

$A \subseteq \mathbf{X}_2$	$m_1^{\downarrow\{2\}}(A)$
$\{b\}$	0.3
$\{\bar{b}\}$	0.5
$\{b, \bar{b}\}$	0.2

$A \subseteq \mathbf{X}_2$	$m_2^{\downarrow\{2\}}(A)$
$\{b\}$	0
$\{\bar{b}\}$	0.5
$\{b, \bar{b}\}$	0.5

TABLE III
BASIC ASSIGNMENT $m_1 \triangleright m_2$ FOR EXAMPLE 1.

	$C \subseteq \mathbf{X}_{\{1,2,3\}}$	$C^{\downarrow\{1,2\}} \otimes C^{\downarrow\{2,3\}}$	$(m_1 \triangleright m_2)(C)$
[a]	$\{a\bar{b}c\}$	$\{a\bar{b}\} \otimes \{\bar{b}c\}$	0.3
[a]	$\{a\bar{b}\bar{c}\}$	$\{a\bar{b}\} \otimes \{\bar{b}\bar{c}\}$	0.1
[a]	$\{a\bar{b}c, a\bar{b}\bar{c}\}$	$\{a\bar{b}\} \otimes \{\bar{b}c, \bar{b}\bar{c}\}$	0.1
[b]	$\{abc, ab\bar{c}\}$	$\{ab\} \otimes \mathbf{X}_1$	0.1
[b]	$\{\bar{a}bc, \bar{a}b\bar{c}\}$	$\{\bar{a}b\} \otimes \mathbf{X}_1$	0.2
[a]	$\{abc, \bar{a}bc, a\bar{b}\bar{c}, \bar{a}\bar{b}\bar{c}\}$	$\mathbf{X}_{\{1,2\}} \otimes \{bc, \bar{b}\bar{c}\}$	0.04
[a]	$\left\{\begin{matrix} abc, \bar{a}bc, a\bar{b}c \\ \bar{a}\bar{b}c, a\bar{b}\bar{c}, \bar{a}\bar{b}\bar{c} \end{matrix}\right\}$	$\mathbf{X}_{\{1,2\}} \otimes \{bc, \bar{b}c, \bar{b}\bar{c}\}$	0.12
[a]	$\left\{\begin{matrix} abc, \bar{a}bc, a\bar{b}c, \bar{a}\bar{b}c \\ ab\bar{c}, \bar{a}b\bar{c}, a\bar{b}\bar{c}, \bar{a}\bar{b}\bar{c} \end{matrix}\right\}$	$\mathbf{X}_{\{1,2\}} \otimes \mathbf{X}_{\{2,3\}}$	0.04

IV. Basic Properties of the Operator of Composition

Let us stress, for the reader familiar with the Dempster's rule of combination [7], that the introduced operator is something quite different.

First, Dempster's rule of combination was defined for two basic assignments defined on the same frame of discernment. In contrast to this, there is no restriction regarding frames of discernments of arguments connected with the introduced operator of composition. Nevertheless, composition of basic assignments defined on the same frame of discernment is uninteresting, because in this case the result is always the first argument - see property (2) of Lemma 1.

Moreover, Dempster's rule of combination (for $C \neq \emptyset$)

$$(m_1 \oplus m_2)(C) = \frac{\sum\limits_{A \cap B = C} m_1(A) \cdot m_2(B)}{1 - \sum\limits_{A \cap B = \emptyset} m_1(A) \cdot m_2(B)}$$

equals $(m_2 \oplus m_1)(C)$; the respective operator $\oplus$ is commutative, which is not the case for the operator $\triangleright$ - see also property (3) of Lemma 1.

The reader should keep in mind that the operator of composition was designed for the situations when one has two basic assignments defined on different frames of discernment and wants to get a new basic assignments defined on a larger frame of discernment incorporating (as much as possible of) the information contained in the original basic assignments.

The following assertion recollects the most important properties of the operator of composition.

Lemma 1: For arbitrary two basic assignments m_1 on $\mathbf{X}_{K_1}$ and m_2 on $\mathbf{X}_{K_2}$ the following properties hold true:

1) $m_1 \triangleright m_2$ *is a basic assignment on* $\mathbf{X}_{K_1 \cup K_2}$.
2) $(m_1 \triangleright m_2)^{\downarrow K_1} = m_1$.
3) $m_1 \triangleright m_2 = m_2 \triangleright m_1 \iff m_1^{\downarrow K_1 \cap K_2} = m_2^{\downarrow K_1 \cap K_2}$.
4) *If* $L \subseteq K_1$ *then* $m_1^{\downarrow L} \triangleright m_1 = m_1$.

Proof: can be found in [5] - Lemma 1. ∎

Realize that property (3) of the preceding Lemma says that the operator is commutative if and only if it is applied to two projective basic assignments. Generally, it is neither commutative nor associative.

V. Multidimensional Models

Consider a sequence of basic assignments $m_1, m_2, \ldots, m_n$ defined on $\mathbf{X}_{K_1}$, $\mathbf{X}_{K_2}, \ldots$, $\mathbf{X}_{K_n}$, respectively. Assume all these assignments are low-dimensional and therefore representable with a reasonable number of parameters. Applying the operator of composition $n-1$ times, one can construct their multiple composition

$$m_1 \triangleright m_2 \triangleright \ldots \triangleright m_n,$$

which may be a multidimensional basic assignment of a rather high dimension. Before starting discussing properties of this expression we have to explain, however, how to understand it; we always apply the operators from left to right:

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

Therefore, to define a multidimensional assignment in this form, it is enough to specify an ordered sequence, we call it a *generating sequence*, of low-dimensional basic assignments. It is obvious that for any permutation $j_1, j_2, \ldots, j_n$ of indices $1, \ldots, n$ the expression

$$\pi_{j_1} \triangleright \pi_{j_2} \triangleright \ldots \triangleright \pi_{j_n}$$

determines a basic assignment with the same frame of discernment, however, for different permutations these basic assignments can differ from one another. In this way, a natural question arises: Which permutation defines a basic assignment with the most advantageous properties? The answer to this question is given in the following definition.

Definition 2: An generating sequence of basic assignments $m_1, m_2, \ldots, m_n$ is said to be *perfect* if

$$\begin{aligned} m_1 \triangleright m_2 &= m_2 \triangleright m_1, \\ (m_1 \triangleright m_2) \triangleright m_3 &= m_3 \triangleright (m_1 \triangleright m_2), \\ &\vdots \\ (m_1 \triangleright \cdots \triangleright m_{n-1}) \triangleright m_n &= m_n \triangleright (m_1 \triangleright \cdots \triangleright m_{n-1}). \end{aligned}$$

The following characterization theorem expresses perhaps the most important result concerning perfect sequences. It says that they compose multidimensional basic assignments that are extensions of all the assignments from which the joint one is composed.

Theorem 1: The sequence $m_1, m_2, \ldots, m_n$ is perfect iff all the basic assignments $m_1, m_2, \ldots, m_n$ are marginal to basic assignment $m_1 \triangleright m_2 \triangleright \ldots \triangleright m_n$.

Proof: can be found in [5] - Theorem 1. ∎

VI. Composition of Bayesian Basic Assignments

It is well known that if all focal elements of a basic assignment m are *singletons*, i.e. if $m(A) > 0$ implies that $|A| = 1$, then this basic assignment corresponds to a probability distribution, and it is why some authors call it *Bayesian basic assignment*. Regarding the fact that operators of composition were originally defined for composition of probability distributions[1] a natural question arises: What is the relation of compositional models in these two theoretical frameworks? To answer this question we shall compare the properties of the corresponding operators of composition. But first, let us recollect how the operator of composition is defined in its probabilistic version.

Let us start considering probability distributions p_i defined on $\mathbf{X}_{K_i}$ (i.e. $p_i : \mathbf{X}_{K_i} \longrightarrow [0,1]$ and $\sum_{x \in \mathbf{X}_{K_i}} p_i(x) = 1$). Analogously to the notation used for basic assignments, their

[1]Probabilistic compositional models were designed as a non-graphical alternative to Bayesian networks and other Graphical Markov models in [4].

marginal distributions (for $L \subset K_i$) will be denoted $p_i^{\downarrow L}$. Realize that $p_i(\emptyset) = 0$, but $p_i^{\downarrow\emptyset}(\emptyset) = 1$.

Definition 3: Consider arbitrary two probability distributions p_1 and p_2 defined on $\mathbf{X}_{K_1}, \mathbf{X}_{K_2}$, respectively ($K_1 \neq \emptyset \neq K_2$). If $p_1^{\downarrow K_1\cap K_2}$ is dominated by $p_2^{\downarrow K_1\cap K_2}$, i.e.

$$\forall z \in \mathbf{X}_{K_1\cap K_2} \quad p_2^{\downarrow K_1\cap K_2}(z) = 0 \Longrightarrow p_1^{\downarrow K_1\cap K_2}(z) = 0,$$

then $p_1 \triangleright p_2$ is for all $x \in \mathbf{X}_{K\cup L}$ defined by the expression

$$(p_1 \triangleright p_2)(x) = \frac{p_1(x^{\downarrow K_1}) \cdot p_2(x^{\downarrow K_2})}{p_2^{\downarrow K_1\cap K_2}(x^{\downarrow K_1\cap K_2})}.$$

(In case of necessity we define $\frac{0\cdot 0}{0} = 0$.) Otherwise the composition $p_1 \triangleright p_2$ remains undefined.

The reader certainly noticed the main difference between the definitions of operators of composition in the two considered theoretical settings: In contrast to composition of basic assignments, it may happen that the composition of probability distributions in not defined. It occurs when $p_2^{\downarrow K_1\cap K_2}$ does not dominate $p_1^{\downarrow K_1\cap K_2}$. In other words, it is undefined if there would be for some $x \in \mathbf{X}_{K\cup L}$ value $(p_1 \triangleright p_2)(x)$ defined by an indeterminate term

$$(p_1 \triangleright p_2)(x) = \frac{p_1(x^{\downarrow K_1}) \cdot 0}{0}$$

with $p_1(x^{\downarrow K_1}) > 0$.

In [5] we proved that if we compose by the operator[2] of composition two Bayesian basic assignments, such that the corresponding probability distributions may be composed by the probabilistic operator of composition (i.e. the probabilistic composition is defined) then the resulting distribution is again Bayesian. The assertion we are about to present here is a little bit stronger: It says that the resulting compositions coincide.

Lemma 2: Let m_1 and m_2 be Bayesian basic assignments on $\mathbf{X}_{K_1}$ and $\mathbf{X}_{K_2}$, respectively, for which

$${m_2}^{\downarrow K_1\cap K_1}(A) = 0 \implies {m_1}^{\downarrow K_1\cap K_2}(A) = 0 \qquad (1)$$

for any $A \subseteq \mathbf{X}_{K_1\cap K_2}$. Let p_1 and p_2 be probabilistic distributions for which

$$\forall x \in \mathbf{X}_{K_1} \quad m_1(\{x\}) = p_1(x);$$
$$\forall y \in \mathbf{X}_{K_2} \quad m_2(\{y\}) = p_2(y).$$

Then $m_1 \triangleright m_2$ is a Bayesian basic assignment and

$$\forall z \in \mathbf{X}_{K_1\cup K_2}(m_1 \triangleright m_2)(\{z\}) = (p_1 \triangleright p_2)(z).$$

Proof: To prove that basic assignment $m_1 \triangleright m_2$ is Bayesian, it is enough to show that if $A \subseteq \mathbf{X}_{K_1\cup K_2}$ is not a singleton then $(m_1 \triangleright m_2)(A) = 0$.

Consider any $A \subseteq \mathbf{X}_{K_1\cup K_2}$ that is not a singleton. Therefore there must exist two two different elements $x, y \in A$. Since $x \neq y$ then either $x^{\downarrow K_1} \neq y^{\downarrow K_1}$ or $x^{\downarrow K_2} \neq y^{\downarrow K_2}$ (or both). Therefore either $A^{\downarrow K_1}$ or $A^{\downarrow K_2}$ is not a singleton and therefore $m_1(A^{\downarrow K_1}) \cdot m_2(A^{\downarrow K_2}) = 0$. This means that if ${m_2}^{\downarrow K_1\cap K_2}(A^{\downarrow K_1\cap K_2}) > 0$ then, due to Definition 1, $(m_1 \triangleright m_2)(A) = 0$.

If ${m_2}^{\downarrow K_1\cap K_2}(A^{\downarrow K_1\cap K_2}) = 0$ then, because we assume the validity of implication (1), ${m_1}^{\downarrow K_1\cap K_2}(A^{\downarrow K_1\cap K_2}) = 0$ and therefore also $m_1(A^{\downarrow K_1}) = 0$. Therefore, according to Definition 1, $(m_1 \triangleright m_2)(A) = 0$, too. So, we have proved that $m_1 \triangleright m_2$ is Bayesian.

Now, consider a singleton $A = \{x\}$ for some $x \in \mathbf{X}_{K_1\cup K_2}$. If $m_2^{\downarrow K_1\cap K_2}(A^{\downarrow K_1\cap K_2}) = p_2(x^{\downarrow K_1\cap K_2}) > 0$, point [a] of Definition 1 yields

$$(m_1 \triangleright m_2)(A) = \frac{m_1(A^{\downarrow K_1}) \cdot m_2(A^{\downarrow K_2})}{m_2^{\downarrow K_1\cap K_2}(A^{\downarrow K_1\cap K_2})},$$

and Definition 3 gives

$$\begin{aligned}(p_1 \triangleright p_2)(x) &= \frac{p_1(x^{\downarrow K_1}) \cdot p_2(x^{\downarrow K_2})}{p_2^{\downarrow K_1\cap K_2}(x^{\downarrow K_1\cap K_2})} \\ &= \frac{m_1(A^{\downarrow K_1}) \cdot m_2(A^{\downarrow K_2})}{m_2^{\downarrow K_1\cap K_2}(A^{\downarrow K_1\cap K_2})} = (m_1 \triangleright m_2)(A).\end{aligned}$$

Similarly, if $m_2^{\downarrow K_1\cap K_2}(A^{\downarrow K_1\cap K_2}) = p_2(x^{\downarrow K_1\cap K_2}) = 0$, we get according to point[3] [c] of Definition 1 that $(m_1 \triangleright m_2)(A) = 0$, and according to Definition 3

$$(p_1 \triangleright p_2)(x) = \frac{p_1(x^{\downarrow K_1}) \cdot p_2(x^{\downarrow K_2})}{p_2^{\downarrow K_1\cap K_2}(x^{\downarrow K_1\cap K_2})} = \frac{0 \cdot 0}{0} = 0,$$

which finishes the proof. ■

It is a direct conclusion following from Lemma 2 that the probabilistic operator of composition meets all the properties presented in Lemma 1. So, the following assertion holds true.

Lemma 3: Consider three probability distributions p_1, p_2, p_3, defined on $\mathbf{X}_{K_1}, \mathbf{X}_{K_2}, \mathbf{X}_{K_3}$, respectively. If $p_1 \triangleright p_2$ is defined then

1) *$p_1 \triangleright p_2$ is a probability distribution on $\mathbf{X}_{K_1\cup K_2}$;*
2) *$(p_1 \triangleright p_2)^{\downarrow K_1} = p_1$;*
3) *$p_1 \triangleright p_2 = p_2 \triangleright p_1 \iff p_1^{\downarrow K_1\cap K_2} = p_2^{\downarrow K_1\cap K_2}$;*
4) *$L \subseteq K_1 \implies p_1^{\downarrow L} \triangleright p_1 = p_1$;*

Let us conclude the section with a simple example (borrowed from [5]) showing a situation when $p_1 \triangleright p_2$ is undefined.

Example 2: Let $\mathbf{X}_1, \mathbf{X}_2$ and $\mathbf{X}_3$ be as in the previous example and consider the following Bayesian basic assignments m_1 and m_2 on $\mathbf{X}_1 \times \mathbf{X}_2$ and $\mathbf{X}_2 \times \mathbf{X}_3$, respectively (realize that since m_1, m_2 are Bayesian, $m_1(A) = m_2(A) = 0$ for all A, for which $|A| > 1$):

$$m_1(\{ab\}) = m_1(\{a\bar{b}\}) = m_1(\{\bar{a}b\}) = m_1(\{\bar{a}\bar{b}\}) = 0.25,$$
$$m_2(\{bc\}) = m_2(\{b\bar{c}\}) = 0.5,$$
$$m_2(\{\bar{b}c\}) = m_2(\{\bar{b}\bar{c}\}) = 0.$$

[2]Notice that by Definitions 1 and 3 we have introduced two operators of composition, both of which are denoted by the same symbol $\triangleright$. We believe that it is obvious that for composition of probability distributions one has to apply the probabilistic version, i.e. Definition 3, whilst for composition of basic assignments one has to apply operator from Definition 1.

[3]Notice that for singleton $A \subseteq \mathbf{X}_{K_1\cup K_2}$, $A = A^{\downarrow K_1} \otimes A^{\downarrow K_2}$ but $A \neq A^{\downarrow K_1} \times \mathbf{X}_{K_2\setminus K_1}$.

TABLE IV

SET OF PROBABILITY DISTRIBUTIONS CORRESPONDING TO $m_1 \triangleright m_2$ FROM EXAMPLE 2.

	a		$\bar{a}$	
	b	$\bar{b}$	b	$\bar{b}$
c	0.125	α	0.125	β
$\bar{c}$	0.125	$1-\alpha$	0.125	$1-\beta$

Let us compute $m_1 \triangleright m_2$ for singletons $\{x_1x_2x_3\} \in \mathbf{X}_1 \times \mathbf{X}_2 \times \mathbf{X}_3$. If $x_2 = b$ then

$$(m_1 \triangleright m_2)(\{x_1bx_3\}) = \frac{m_1(\{x_1b\}) \cdot m_2(\{bx_3\})}{m_2^{\downarrow\{2\}}(\{b\})} = \frac{0.25 \cdot 0.5}{1} = 0.125,$$

and for singletons $\{x_1\bar{b}x_3\}$ we get

$$(m_1 \triangleright m_2)(\{x_1\bar{b}x_3\}) = 0,$$

because $m_2^{\downarrow\{2\}}(\{\bar{b}\}) = 0$. In this case, however, we get (according to the point [b] of Definition 1)

$$(m_1 \triangleright m_2)(\{x_1\bar{b}\} \times \mathbf{X}_3) = m_1(\{x_1\bar{b}\}) = 0.25.$$

This means that in this case there are 6 focal elements of $m_1 \triangleright m_2$, namely 4 singletons:

$$\{x_1bx_3\}, \text{ for } x_1 \in \mathbf{X}_1, x_3 \in \mathbf{X}_3,$$

for which $(m_1 \triangleright m_2)(\{x_1bx_3\}) = 0.125$, and 2 two-element sets:

$$\{x_1\bar{b}\} \times \mathbf{X}_3 = \{x_1\bar{b}c, x_1\bar{b}\bar{c}\}, \text{ for } x_1 \in \mathbf{X}_1,$$

for which $(m_1 \triangleright m_2)(\{x_1\bar{b}\} \times \mathbf{X}_3) = 0.25$.

Regarding this example, let us remark two points:

1) $m_1 \triangleright m_2$ corresponds to a whole set of probability distributions, which are recorded in Table IV (for any $\alpha \in [0, 0.25]$ and $\beta \in [0, 0.25]$). It is in correspondence with a general rule holding for any two Bayesian basic assignments m_1 and m_2 and the corresponding probability distributions p_1 and p_2: If $p_1 \triangleright p_2$ is not defined then $m_1 \triangleright m_2$ is not Bayesian.
2) In contrast to $m_1 \triangleright m_2$, $m_2 \triangleright m_1$ is a Bayesian basic assignment. This basic assignment has 4 focal elements:

$$(m_2 \triangleright m_1)(\{x_1bx_3\}) = 0.25, \text{ for } x_1 \in \mathbf{X}_1, x_3 \in \mathbf{X}_3.$$

VII. GENERALIZATION OF PROBABILISTIC MODELS

In this section we shall make a couple of suggestions enabling us to understand multidimensional models of basic assignments as a real enrichment of probabilistic models. First let us have a look how the concept of conditional independence was introduced in these two theoretical settings.

Consider three disjoint sets $I, J, K \subset N$ ($I \neq \emptyset \neq J$) and a probability distribution p on $\mathbf{X}_N$. We say that for distribution p groups of variables X_I and X_J are *conditionally independent given variables* X_K if for all $x \in \mathbf{X}_{I\cup J\cup K}$ the following equality holds true

$$p^{\downarrow I\cup J\cup K}(x) \cdot p^{\downarrow K}(x^{\downarrow K}) = p^{\downarrow I\cup K}(x^{\downarrow I\cup K}) \cdot p^{\downarrow J\cup K}(x^{\downarrow J\cup K}).$$

It is well known that this is equivalent to the fact that

$$p^{\downarrow I\cup J\cup K}(x) = p^{\downarrow I\cup K}(x^{\downarrow I\cup K}) \cdot p^{\downarrow J\cup K}(x^{\downarrow J}|x^{\downarrow K}),$$

or, using the probabilistic operator of composition

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

How is it for basic assignments? Answering the question is not so easy because of the fact that this notion for belief functions was introduced in several different ways. Perhaps the most frequent (and maybe also with the greatest number of supporters) is the one, which can be easily defined with the help of *commonality function*. Using notation of Studený [9], commonality function Com_m is defined for basic assignment m (assuming that m is defined on $\mathbf{X}_N$) for each $A \subset \mathbf{X}_N$ by a simple formula

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

Ben Yaghlane *et al.* [1], [2], [3] define the concept of conditional non-interactivity (as well as Shenoy defines his concept of conditional independence [8]) in the way that variables X_I and variables X_J are *conditionally non-interactive* given variables X_K if and only if for all $A \subseteq \mathbf{X}_N$

$$Com_{m^{\downarrow I\cup J\cup K}}(A^{\downarrow I\cup J\cup K}) \cdot Com_{m^{\downarrow K}}(A^{\downarrow K}) = Com_{m^{\downarrow I\cup K}}(A^{\downarrow I\cup K}) \cdot Com_{m^{\downarrow J\cup K}}(A^{\downarrow J\cup K}).$$

In this paper we shall denote this property by

$$X_I \perp\!\!\!\perp_{[m]} X_J | X_K.$$

Unfortunately, for basic assignments it does not hold true that $X_I \perp\!\!\!\perp_{[m]} X_J|X_K$ if and only if the basic marginal assignment $m^{\downarrow I\cup J\cup K}$ *factorizes* in the following sense

$$m^{\downarrow I\cup J\cup K} = m^{\downarrow I\cup K} \triangleright m^{\downarrow J\cup K}. \quad (2)$$

Nevertheless, there are still properties indicating a similarity of these two notions. First, Ben Yaghlane *et al.* in [2] showed that if $X_I \perp\!\!\!\perp_{[m]} X_J|X_K$ then all focal elements of $m^{\downarrow I\cup J\cup K}$ (i.e. sets $A \subseteq \mathbf{X}_{I\cup J\cup K}$, for which $m^{\downarrow I\cup J\cup K}(A) > 0$) are Z-layered rectangles, which are nothing else, as we said before, than sets $A \subseteq \mathbf{X}_{I\cup J\cup K}$, which can be expressed as an extension of its respective projections:

$$A = A^{\downarrow I\cup K} \otimes A^{\downarrow J\cup K}.$$

Therefore, combining the mentioned Ben Yaghlane *et al.* property with Definition 1 we get the following simple assertion.

Assertion: Consider a basic assignment m on $\mathbf{X}_N$ and three disjoint subsets $I, J, K \subset N$ ($I \neq \emptyset \neq J$). If $A \subseteq \mathbf{X}_{I\cup J\cup K}$ is a focal element of $m^{\downarrow I\cup J\cup K}$ and $A \neq A^{\downarrow I\cup K} \otimes A^{\downarrow J\cup K}$ then neither of the following two expressions holds true:

$$X_I \perp\!\!\!\perp_{[m]} X_J|X_K,$$

and

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

So, the first property connecting the concepts of conditional non-interactivity and factorization for basic assignments is that any of them guarantees that the focal elements of the respective basic assignment can be expressed as an extension of its corresponding projections (Z-layered rectangles in the language of Ben Yaghlane *et al.*).

Another connecting property says that these notions coincide for Bayesian basic assignments. Namely, in [9] Studený claims that for Bayesian basic assignments the concept of conditional non-interactivity coincides with the concept of conditional independence of the corresponding probability distribution. Due to Lemma 2 the same holds also for the concept of factorization in the sense of equation (2).

Let us now pinpoint the difference between the studied concepts. In [2] the authors admit that their concept of conditional non-interactivity (as showed by Studený) is *not consistent with marginalization*[10], [11]. This means that it may happen that there are two consistent basic assignments m_1 and m_2 defined on $\mathbf{X}_{I\cup K}$ and $\mathbf{X}_{J\cup K}$, respectively (I, J, K disjoint, $I \neq \emptyset \neq J$), for which there does not exist a basic assignment m on $\mathbf{X}_{I\cup J\cup K}$, such that m_1 and m_2 would be its marginal assignments and simultaneously $X_I \perp\!\!\!\perp_{[m]} X_J | X_K$. For an example see [2]. Such a situation, however, cannot happen for the concept of factorization, since $m_1 \triangleright m_2$ is always defined and $(m_1 \triangleright m_2)^{\downarrow K_1} \triangleright (m_1 \triangleright m_2)^{\downarrow K_2} = m_1 \triangleright m_2$.

Taking into account also the fact that, as we showed in [6], factorization in the sense of equality (2) meets all the semigraphoid axioms, we are making the following suggestion.

PROPOSAL 1: Introduce the concept of conditional independence relation for basic assignments with the help of factorization in the sense of equality (2).

Probabilistic compositional models have, from the point of view of practical applications, a disadvantage that a necessary composition need not be defined. It is true that it may happen only in situations when one composes probability distributions which are not consistent. But it may easily occur when one constructs a model from data from different sources or when a source with missing data is considered. To avoid this problem we propose the following solution.

PROPOSAL 2: Apply the operator of composition designed for basic assignments (Definition 1) even when handling probability distributions and consider in some cases sets of probability distributions.

Surprisingly enough, realization of this proposal need not increase computational complexity of the used algorithms. This statement is based on the fact that space complexity of these models is not higher that space complexity of the corresponding probabilistic models. Going back to the Example 2 we can see that basic assignment $m_1 \triangleright m_2$ has a smaller number of focal elements (i.e. it is defined with a smaller number of parameters) than a general 3-dimensional probability distribution. In fact, it is a rule, that when composing Bayesian basic assignments then the resulting basic assignment does not have more focal elements than the number of points of the corresponding frame of discernment. This statement is trivial when the result is again Bayesian assignment, however it is important to realize it holds also when the result is non-Bayesian.

VIII. Conclusion

In the paper we have introduced an operator of composition for basic assignments, which enables us to construct multidimensional models from a sequences of low-dimensional assignments. We showed that these models are true generalization of probabilistic models and therefore we propose to use them whenever classical probabilistic model, due to incoherence of low-dimensional probability distributions, does not exist.

To increase consistency of probabilistic models and a wider class of models constructed from basic assignments, we proposed also to introduce a new concept of conditional independence for basic assignments: the concept corresponding to factorization.

Acknowledgment

The research was financially supported by Ministry of Education of the Czech Republic by grants no. 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] E. Ben Yaghlane, Ph. Smets, and K. Mellouli, “Directed Evidential Networks with Conditional Belief functions,” *Proc. of ECSQARU 2003*, LNAI 2711, Springer, pp. 291–305, 2003.

[4] 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.

[5] R. Jiroušek, J. Vejnarová and M. Daniel, ”Compositional models of belief functions,” *Proc. of the 5th Symposium on Imprecise Probabilities and Their Applications* (G. de Cooman, J. Vejnarová and M. Zaffalon, eds.), Charles University Press, Praha, pp. 243–252, 2007.

[6] R. Jiroušek ”Conditional Independence and Factorization of Multidimensional Models,” accepted for publications in: *Proc. of the 2008 IEEE World Congress on Computational Intelligence (WCCI 2008)*, HongKong.

[7] G. Shafer, *A mathematical theory of evidence.*, Princeton University Press, 1976.

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

[9] M. Studený, “Formal properties of conditional independence in different calculi of AI,” *Proceedings of European Conference on Symbolic and quantitative Approaches to Reasoning and Uncertainty ECSQARU’93*, (K. Clarke, R. Kruse and S. Moral, eds.). Springer-Verlag, 1993, pp. 341–351.

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

[11] M. Studený, *Probabilistic Conditional Independence Structures.* Springer, London, 2005.