Appeared in: Belief Functions: Theory and Applications, T. Denoeux and M.-H. Masson (eds.),
Advances in Intelligent and Soft Computing 164, 2012, pp. 221--228, Springer-Verlag, Berlin.

Compositional Models in Valuation-Based Systems

Radim Jiroušek and Prakash P. Shenoy

Abstract Compositional models were initially described for discrete probability theory, and later extended for possibility theory, and Dempster-Shafer (D-S) theory of evidence. Valuation-based systems (VBS) can be considered as a generic uncertainty framework that has many uncertainty calculi, such as probability theory, a version of possibility theory where combination is the product t-norm, Spohn's epistemic belief theory, and D-S belief function theory, as special cases. In this paper, we describe compositional models for the VBS framework using the semantics of no-double counting. We show that the compositional model defined here for belief functions differs from the one studied by Jiroušek, Vejnarová, and Daniel. The latter model can be described in the VBS framework, but with a combination operation that is different from Dempster's rule.

1 Introduction

Compositional models were initially described for discrete probability theory [4, 5]. They were later extended by Vejnarová [14] for possibility theory, and in [6] for belief functions in the Dempster-Shafer (D-S) belief function theory. In this paper, we use the valuation-based systems (VBS) framework [10] to extend compositional models to all uncertainty calculi captured by the VBS framework, which includes calculi such as probability theory, a version of possibility theory with the product t-norm, Spohn's epistemic belief theory, and D-S belief function theory.

We start by recalling the necessary basic notions of the VBS framework (most of the material is taken from [10]).

Radim Jiroušek
Faculty of Management, University of Economics, Jindřichův Hradec and Prague, Czech Republic, e-mail: radim@utia.cas.cz

Prakash P. Shenoy
School of Business, University of Kansas, Lawrence, KS, USA e-mail: pshenoy@ku.edu

2 Valuation-Based Systems

VBS consists of two parts — a static part that is concerned with representation of knowledge, and a dynamic part that is concerned with reasoning.

The static part consists of objects called variables and valuations. Let Φ denote a finite set whose elements are called *variables*. Elements of Φ are denoted by upper-case Roman alphabets such as X, Y, Z, etc. Subsets of Φ are denoted by lower-case Roman alphabets such as r, s, t, etc.

Let Ψ denote a finite set whose elements are called *valuations*. Elements of Ψ are denoted by lower-case Greek alphabets such as ρ, σ, τ, etc. Each valuation is associated with a subset of variables, and represents some knowledge about the variables in the subset. Thus, we say that ρ is a valuation for r, where $r \subseteq \Phi$ is the subset associated with ρ.

We identify a subset of valuations $\Psi_n \subset \Psi$, whose elements are called *normal* valuations. Normal valuations are valuations that are coherent in some sense. In D-S belief function theory, normal valuations are basic probability assignment potentials whose values for all non-empty subsets add to one.

The dynamic part of VBS consists of three operators — combination, marginalization, and removal — that are used to make inferences from the knowledge encoded in a VBS. We define these operators using axioms.

Combination. The first operator is the *combination* operator $\oplus\colon \Psi \times \Psi \to \Psi_n$, which represents aggregation of knowledge. It must satisfy the following three axioms:

1. (*Domain*) If ρ is a valuation for r, and σ is a valuation for s, then $\rho \oplus \sigma$ is a normal valuation for $r \cup s$.
2. (*Commutativity*) $\rho \oplus \sigma = \sigma \oplus \rho$.
3. (*Associativity*) $\rho \oplus (\sigma \oplus \tau) = (\rho \oplus \sigma) \oplus \tau$.

The domain axiom expresses the fact that if ρ represents some knowledge about variables in r, and σ represents some knowledge about variables in s, then $\rho \oplus \sigma$ represents the aggregated knowledge about variables in $r \cup s$. The commutativity and associativity axioms reflect the fact that the sequence in which knowledge is aggregated makes no difference in the aggregated result.

The set of all normal valuations with the combination operator $\oplus$ forms a commutative semigroup. We let $\iota_\emptyset$ denote the (unique) identity valuation of this semigroup. Thus, for any normal valuation ρ, $\rho \oplus \iota_\emptyset = \rho$.

The set of all normal valuations for $s \subseteq \Phi$ with the combination operator $\oplus$ also forms a commutative semigroup (which is different from the semigroup discussed in the previous paragraph). Let ι_s denote the (unique) identity for this semigroup. Thus, for any normal valuation σ for s, $\sigma \oplus \iota_s = \sigma$.

Notice that in general $\rho \oplus \rho \neq \rho$. Thus, it is important to ensure that we do not double count knowledge when double counting matters, i.e., it is okay to double count knowledge ρ that is idempotent, i.e., $\rho \oplus \rho = \rho$. In representing our knowledge as valuations in Ψ, we have to ensure that there is no double counting of non-idempotent knowledge.

Marginalization. Another operator is marginalization $-X : \Psi \to \Psi$, which allows us to coarsen knowledge by marginalizing X out of the domain of a valuation. It must satisfy the following four axioms:

1. (*Domain*) If ρ is a valuation for r, and $X \in r$, then ρ^{-X} is a valuation for $r \setminus \{X\}$.
2. (*Normal*) ρ^{-X} is normal if and only if ρ is normal.
3. (*Order does not matter*) If ρ is a valuation for r, $X \in r$, and $Y \in r$, then $(\rho^{-X})^{-Y} = (\rho^{-Y})^{-X}$, which is denoted by $\rho^{-\{X,Y\}}$.
4. (*Local computation*) If ρ and σ are valuations for r and s, respectively, $X \in r$, and $X \notin s$, then $(\rho \oplus \sigma)^{-X} = (\rho^{-X}) \oplus \sigma$.

The domain axiom is self-explanatory. Marginalization preserves normal (and non-normal) property of valuations. The order does not matter axiom dictates that when we coarsen knowledge by marginalizing out several variables, the order in which the variables are marginalized does not matter in the final result. Occasionally, we let $\rho^{\downarrow r \setminus \{X,Y\}}$ denote $\rho^{-\{X,Y\}}$.

Removal. The removal operator $\ominus : \Psi \times \Psi_n \to \Psi_n$ represents removing knowledge in the second valuation from the knowledge in the first valuation. It must satisfy the following three axioms:

1. (*Domain*): Suppose σ is a valuation for s and ρ is a normal valuation for r. Then $\sigma \ominus \rho$ is a normal valuation for $r \cup s$.
2. (*Identity*): For each normal valuation ρ for r, $\rho \oplus \rho \ominus \rho = \rho$. Thus, $\rho \ominus \rho$ acts as an identity for ρ, and we denote $\rho \ominus \rho$ by ι_ρ. Thus, $\rho \oplus \iota_\rho = \rho$.
3. (*Combination and Removal*): Suppose π and θ are valuations, and suppose ρ is a normal valuation. Then, $(\pi \oplus \theta) \ominus \rho = \pi \oplus (\theta \ominus \rho)$.

We call $\sigma \ominus \rho$ the valuation resulting after removing ρ from σ. The identity axiom defines the removal operator as an inverse of the combination operator.

In [10], a number of properties of combination, marginalization, and removal operators are proved. For example, suppose π, σ, θ are valuations for p, s, and t, respectively, ρ is a normal valuation for r, $X \in s$, and $X \notin r$. Then, $(\pi \oplus \theta) \ominus \rho = (\pi \ominus \rho) \oplus \theta$, and $(\sigma \ominus \rho)^{-X} = \sigma^{-X} \ominus \rho$.

3 VBS for D-S Belief Function Theory

In D-S belief function theory, we can use either basic probability assignments, or belief functions, or plausibility functions, or commonality functions, to represent knowledge. Here, we use only basic probability assignments.

Basic Probability Assignment. A *basic probability assignment* (bpa) μ for s is a function $\mu : 2^{\Omega_s} \to \mathbb{R}$ such that $\mu(\mathbf{a}) \geq 0$ for all $\mathbf{a} \in 2^{\Omega_s}$, and $\sum\{\mu(\mathbf{a}) \mid \mathbf{a} \in 2^{\Omega_s}\} = 1$.

B-Valuations. A b-valuation σ for s is a function $\sigma : 2^{\Omega_s} \to \mathbb{R}$. We say σ is *normal* if $\sum\{\sigma(\mathbf{a}) \mid \mathbf{a} \in 2^{\Omega_s}\} = 1$, and we say σ is *proper* if $\sigma(\mathbf{a}) \geq 0$ for all $\mathbf{a} \in 2^{\Omega_s}$. Proper normal b-valuations represent bpa functions. Normal b-valuations that are not proper are called *pseudo-bpa* .

Set Operations. Suppose r, s, and t are sets of variables, $r \subseteq s$. For $x \in \Omega_s$, $x^{\downarrow r}$ denotes the projection of x into Ω_r. Similarly, for $\mathbf{a} \in 2^{\Omega_s}$, the projection of $\mathbf{a}$ to r, denoted by $\mathbf{a}^{\downarrow r}$, is given by $\mathbf{a}^{\downarrow r} = \{x^{\downarrow r} \mid x \in \mathbf{a}\}$. Also, if $\mathbf{a} \subseteq \Omega_s$, and $\mathbf{b} \subseteq \Omega_t$, then the join of $\mathbf{a}$ and $\mathbf{b}$, denoted by $\mathbf{a} \bowtie \mathbf{b}$ is given by:

$$\mathbf{a} \bowtie \mathbf{b} = \{x \in \Omega_{s \cup t} \mid x^{\downarrow s} \in \mathbf{a}, x^{\downarrow t} \in \mathbf{b}\}. \tag{1}$$

Combination. Suppose ρ and σ are b-valuations for r and s, respectively. Let K denote $\sum\{\rho(\mathbf{b}) \cdot \sigma(\mathbf{c}) \mid \mathbf{b} \subseteq \Omega_r, \mathbf{c} \subseteq \Omega_s \text{ s.t. } \mathbf{b} \bowtie \mathbf{c} = \emptyset\}$. The combination $\rho \oplus \sigma$ is a normal b-valuation for $r \cup s$ given for all $\mathbf{a} \subseteq \Omega_{r \cup s}$ by

$$(\rho \oplus \sigma)(\mathbf{a}) = \begin{cases} K^{-1} \sum\{\rho(\mathbf{b}) \cdot \sigma(\mathbf{c}) \mid \mathbf{b} \subseteq \Omega_r, \mathbf{c} \subseteq \Omega_s \text{ s.t. } \mathbf{b} \bowtie \mathbf{c} = \mathbf{a}\} & \text{if } K \neq 0 \\ 0 & \text{if } K = 0. \end{cases} \tag{2}$$

If $K \neq 0$, then K is a the normalization constant that ensures that $\rho \oplus \sigma$ is a normal b-valuation. It is evident that if ρ and σ are bpa's (proper normal b-valuations), and $K \neq 0$, then $\rho \oplus \sigma$ is a bpa. It can be shown that the definition of combination in Equation (2) satisfies the three axioms of combination.

Marginalization. Suppose σ is a b-valuation for s, and suppose $X \in s$. The marginal σ^{-X} is a b-valuation for $s \setminus \{X\}$ given by

$$\sigma^{-X}(\mathbf{a}) = \sum\{\sigma(\mathbf{b}) \mid \mathbf{b} \in 2^{\Omega_s} \text{ s.t. } \mathbf{b}^{\downarrow s \setminus \{X\}} = \mathbf{a}\} \quad \text{for all } \mathbf{a} \in 2^{\Omega_{s \setminus \{X\}}}. \tag{3}$$

It can be shown that the definition of marginalization in Equation (3) satisfies the four axioms of marginalization.

Removal. Removal is inverse of combination. It is not easy to define removal in terms of b-valuations. For readers familiar with commonality functions, $\oplus$ reduces to pointwise multiplication of commonality functions followed by normalization. Thus, $\sigma \ominus \rho$ is pointwise division of commonality functions corresponding to σ and ρ, followed by normalization. It can be shown that this definition satisfies the three axioms of removal.

Notice that if σ and ρ are proper b-valuations, it is possible that $\sigma \ominus \rho$ is a pseudo-bpa. This may be true even if $r \subseteq s$ and ρ is a marginal of σ.

Convention. For the sake of simplicity, in the rest of this paper we assume that whenever the operator $\oplus$ or $\ominus$ is applied, then the result does not result in the zero valuation, a valuation whose values are identically 0.

4 Compositional Models in VBS

Suppose we have marginals for two overlapping subsets of variables, say for $\{D, G\}$ and $\{D, B\}$. How do we construct a joint distribution for $\{D, G, B\}$ that is consistent with the two marginals (assuming that it exists)? In [4], the operation of "composing" the two marginals to obtain a joint distribution is introduced. One way to view the composition operator is in terms of no double counting. Notice that the two marginals are not distinct since the knowledge of $\{D\}$ is included in both marginals. So, the composition operator should aggregate the knowledge in the two marginals while adjusting for the double counting of knowledge of $\{D\}$.

In practice, it is extremely unlikely we would find marginals on non-disjoint subsets of variables with common marginals. In this case, there does not exist a joint that agrees with both marginals. So we relax the requirements so that the joint distribution that is constructed is required to agree only with the first marginal.

Composition. A general definition of composition is as follows. Suppose ρ and σ are normal valuations for r and s, respectively. The composition of ρ and σ, written as $\rho \triangleright \sigma$, is defined as follows:

$$\rho \triangleright \sigma = \rho \oplus \sigma \ominus \sigma^{\downarrow r \cap s} \tag{4}$$

It can be seen directly from the definition in Equation (4) that the composition operator is, in general, neither commutative nor associative. Its most important properties are summarized in the following lemma.

Lemma. *Suppose ρ and σ are normal valuations for r and s, respectively. Then the following statements hold.*

1. Domain: $\rho \triangleright \sigma$ *is a normal valuation for* $r \cup s$.
2. Composition preserves first marginal: $(\rho \triangleright \sigma)^{\downarrow r} = \rho$.
3. Commutativity under consistency: *If ρ and σ have a common marginal for $r \cap s$, i.e., $\rho^{\downarrow r \cap s} = \sigma^{\downarrow r \cap s}$, then $\rho \triangleright \sigma = \sigma \triangleright \rho$.*
4. Associativity under a special condition: *Suppose τ is a normal valuation for t, and suppose $s \supset (r \cap t)$. Then, $(\rho \triangleright \sigma) \triangleright \tau = \rho \triangleright (\sigma \triangleright \tau)$.*
5. Composition of marginals: *Suppose t is such that $(r \cap s) \subseteq t \subseteq s$. Then*

$$(\rho \triangleright \sigma^{\downarrow t}) \triangleright \sigma = \rho \triangleright \sigma.$$

5 Comparison with an Alternative Compositional Model

For belief functions in the D-S theory, the operator of composition was originally introduced in [6]. Since, as it will be shown in a simple example, it differs from the operator introduced here in Equation (4), we will use for the original operator a slightly different symbol.

Definition Suppose ρ and σ are normal b-valuations for r and s, respectively. The old-composition of ρ and σ, written here as $\rho \trianglerighteq \sigma$, is defined for each $\mathbf{a} \subseteq \Omega_{r \cup s}$ by one of the following expressions:

[1] if $\sigma^{\downarrow r \cap s}(\mathbf{a}^{\downarrow r \cap s}) > 0$ and $\mathbf{a} = \mathbf{a}^{\downarrow r} \bowtie \mathbf{a}^{\downarrow s}$ then $(\rho \trianglerighteq \sigma)(\mathbf{a}) = \frac{\rho(\mathbf{a}^{\downarrow r}) \cdot \sigma(\mathbf{a}^{\downarrow s})}{\sigma^{\downarrow r \cap s}(\mathbf{a}^{\downarrow r \cap s})}$;

[2] if $\sigma^{\downarrow r \cap s}(\mathbf{a}^{\downarrow r \cap s}) = 0$ and $\mathbf{a} = \mathbf{a}^{\downarrow r} \times \Omega_{s \setminus r}$ then $(\rho \trianglerighteq \sigma)(\mathbf{a}) = \rho(\mathbf{a}^{\downarrow r})$;

[3] in all other cases $(\rho \trianglerighteq \sigma)(\mathbf{a}) = 0$.

Example. Consider the Studený's example [1]. Suppose X,Y and Z are variables with state spaces $\Omega_X = \{x,\bar{x}\}$, $\Omega_Y = \{y,\bar{y}\}$, and $\Omega_Z = \{z,\bar{z}\}$. Consider two b-valuations ρ and σ for $\{X,Z\}$ and $\{Y,Z\}$, respectively, each having only two non-zero values: $\rho(\{x\bar{z},\bar{x}z\}) = \rho(\{x\bar{z},\bar{x}\bar{z}\}) = 0.5$ and $\sigma(\{y\bar{z},\bar{y}z\}) = \sigma(\{y\bar{z},\bar{y}\bar{z}\}) = 0.5$.

In [7], it is shown that $\rho \trianglerighteq \sigma$ has also only two non-zero values: $(\rho \trianglerighteq \sigma)(\{xy\bar{z},\bar{x}\bar{y}z\}) = (\rho \trianglerighteq \sigma)(\{xy\bar{z},x\bar{y}\bar{z},\bar{x}y\bar{z},\bar{x}\bar{y}\bar{z}\}) = 0.5$. Thus, we see that $\rho \trianglerighteq \sigma$ is a proper normal b-valuation.

Also, $\rho \oplus \sigma$ is a normal b-valuation with value 0.25 for the following four sets: $\{xy\bar{z},x\bar{y}\bar{z}\},\{xy\bar{z},\bar{x}y\bar{z}\},\{xy\bar{z},\bar{x}\bar{y}z\},\{xy\bar{z},x\bar{y}\bar{z},\bar{x}y\bar{z},\bar{x}\bar{y}\bar{z}\}$. In contrast, $\rho \triangleright \sigma = \rho \oplus \sigma \ominus \sigma^{-Y}$ is a pseudo-bpa since $(\rho \triangleright \sigma)(\{\bar{x}\bar{y}z\}) = -0.25$ (the following are the remaining non-zero values of $\rho \triangleright \sigma$: $(\rho \triangleright \sigma)(\{xy\bar{z},x\bar{y}\bar{z}\}) = 0.25$, $(\rho \triangleright \sigma)(\{xy\bar{z},\bar{x}y\bar{z}\}) = 0.25$, $(\rho \triangleright \sigma)(\{xy\bar{z},\bar{x}\bar{y}z\}) = 0.5$, $(\rho \triangleright \sigma)(\{xy\bar{z},x\bar{y}\bar{z},\bar{x}y\bar{z},\bar{x}\bar{y}\bar{z}\}) = 0.25$).

It is worth mentioning that the same result as $\rho \trianglerighteq \sigma$ is obtained also by the Srivastava-Cogger algorithm [13], but it need not be the case for different values of the ρ and σ b-valuations in this example.

To understand the differences between the two operators of composition, recall that a close connection exists between the combination operator $\oplus$ and a notion of independence. Namely, after combining ρ for X and σ for Y, we get the valuation $\rho \oplus \sigma$ for $\{X,Y\}$, with respect to which variables X and Y are independent. Similarly, if ρ is a valuation for $\{X,Z\}$, and σ is a valuation for $\{Y,Z\}$, with respect to the valuation $\rho \oplus \sigma$ for $\{X,Y,Z\}$, variables X and Y are conditionally independent given Z. However, several other concepts of independence and conditional independence for belief functions exists in the literature. For a non-exhaustive survey, see [1, 2].

In their seminal papers, Dempster [3] and Walley and Fine [15] considered a type of independence that hold for variables X and Y with respect to bpa μ for $\{X,Y\}$ if

$$\mu(\mathbf{a}) = \begin{cases} \mu^{\downarrow X}(\mathbf{a}^{\downarrow X}) \cdot \mu^{\downarrow Y}(\mathbf{a}^{\downarrow Y}) & \text{if } \mathbf{a} = \mathbf{a}^{\downarrow X} \times \mathbf{a}^{\downarrow Y} \\ 0 & \text{otherwise} \end{cases} \quad \text{for all } \mathbf{a} \in \Omega_{\{X,Y\}}. \tag{5}$$

Generalizing this idea, we define an alternative operation of combination, denoted by $\underline{\oplus}$, for b-valuations ρ and σ (for r and s, respectively) as follows. Suppose K denotes $\sum\{\rho(\mathbf{a}^{\downarrow r}) \cdot \sigma(\mathbf{a}^{\downarrow s}) \mid \mathbf{a} \in \Omega_{r\cup s} \text{ s.t. } \mathbf{a} = \mathbf{a}^{\downarrow r} \bowtie \mathbf{a}^{\downarrow s}\}$. The combination $\rho \underline{\oplus} \sigma$ is the b-valuation for $r \cup s$ given for all $\mathbf{a} \in \Omega_{r\cup s}$ by

$$(\rho \underline{\oplus} \sigma)(\mathbf{a}) = \begin{cases} K^{-1} \rho(\mathbf{a}^{\downarrow r})\, \sigma(\mathbf{a}^{\downarrow s}) & \text{if } K > 0,\, \mathbf{a} = \mathbf{a}^{\downarrow r} \bowtie \mathbf{a}^{\downarrow s} \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

It is obvious that $\rho \underline{\oplus} \sigma$ defined in Equation (6) is a proper normal b-valuation for $r \cup s$, and that $\underline{\oplus}$ satisfies all the three axioms of combination.

In a similar way, we define an alternative removal operator $\underline{\ominus}$. Suppose ρ and σ are b-valuations for r and s, respectively, and suppose that ρ is normal. Let K denote $\sum\{\frac{\sigma(\mathbf{a}^{\downarrow s})}{\rho(\mathbf{a}^{\downarrow r})} \mid \mathbf{a} \in \Omega_{r\cup s} \text{ s.t. } \mathbf{a} = \mathbf{a}^{\downarrow r} \bowtie \mathbf{a}^{\downarrow s}, \rho(\mathbf{a}^{\downarrow r}) > 0\}$. $\sigma \underline{\ominus} \rho$ is the b-valuation for $s \cup r$ given for all $\mathbf{a} \in \Omega_{s\cup r}$ by

$$(\sigma \underline{\oplus} \rho)(\mathbf{a}) = \begin{cases} K^{-1}\left(\frac{\sigma(\mathbf{a}^{\downarrow s})}{\rho(\mathbf{a}^{\downarrow r})}\right) & \text{if } K > 0,\, \mathbf{a} = \mathbf{a}^{\downarrow r} \bowtie \mathbf{a}^{\downarrow s},\, \rho(\mathbf{a}^{\downarrow r}) > 0 \\ 0 & \text{otherwise.} \end{cases} \tag{7}$$

Thus, together with marginalization defined as in Section 3, we get an alternative VBS for belief functions in the D-S theory. Let two normal b-valuations ρ and σ for r and s, respectively, be such that

$$\sigma^{\downarrow r\cap s}(\mathbf{x}) = 0 \Longrightarrow \rho^{\downarrow r\cap s}(\mathbf{x}) = 0.$$

Consider $\mathbf{a} \subseteq \Omega_{r\cup s}$ for which $\mathbf{a} = \mathbf{a}^{\downarrow r} \bowtie \mathbf{a}^{\downarrow s}$. Then,

$$(\rho \underline{\oplus} \sigma \underline{\ominus} \sigma^{\downarrow r\cap s})(\mathbf{a}) = \begin{cases} k\left(\frac{\rho(\mathbf{a}^{\downarrow r})\,\sigma(\mathbf{a}^{\downarrow s})}{\sigma^{\downarrow r\cap s}(\mathbf{a}^{\downarrow r\cap s})}\right) & \text{if } \sigma^{\downarrow r\cap s}(\mathbf{a}^{\downarrow r\cap s}) > 0 \\ 0 & \text{otherwise,} \end{cases} \tag{8}$$

which, due to the definition of old-composition, can be rewritten as

$$(\rho \underline{\oplus} \sigma \underline{\ominus} \sigma^{\downarrow r\cap s})(\mathbf{a}) = k\,(\rho \trianglerighteq \sigma)(\mathbf{a}).$$

Notice that because of the above assumption, when computing $\rho \trianglerighteq \sigma$, whenever case [2] of the definition of old composition applies, the value $\rho(\mathbf{a}^{\downarrow r}) = 0$.

Since for all $\mathbf{a} \neq \mathbf{a}^{\downarrow r} \bowtie \mathbf{a}^{\downarrow s}$, $(\rho \underline{\oplus} \sigma \underline{\ominus} \sigma^{\downarrow r\cap s})(\mathbf{a}) = (\rho \trianglerighteq \sigma)(\mathbf{a}) = 0$, we get

$$(\rho \underline{\oplus} \sigma \underline{\ominus} \sigma^{\downarrow r\cap s})(\mathbf{a}) = k\,(\rho \trianglerighteq \sigma)(\mathbf{a}), \quad \text{for all } \mathbf{a} \subseteq \Omega_{r\cup s}.$$

Since we know that both $\rho \underline{\oplus} \sigma \underline{\ominus} \sigma^{\downarrow r\cap s}$ and $\rho \trianglerighteq \sigma$ are normal b-valuations (for the former, it follows from the lemma presented in Section 4; for the latter, it is proved in [6]), it follows that $k = 1$.

Thus, we have shown that the operator of composition defined in [6] can be considered as a special case of composition in a VBS where combination is $\underline{\oplus}$, removal is $\underline{\ominus}$, and marginalization is the same as in the D-S theory.

6 Summary and Conclusions

We have described the VBS framework in general, and described the composition model in the VBS framework using the semantics of no double counting of knowledge. We have compared the compositional model defined in this paper for D-S belief function *theory* with the one described in [6] for belief functions. Our conclusion is that although both of these compositional models are defined for belief functions and its alternative representations (bpa, commonality, etc.), the former is defined for *the* D-S belief function theory (that necessarily entails Dempster's rule of combination), and the latter for a belief function theory that has $\underline{\oplus}$ as the rule of combination. Both of these theories fit in the VBS framework, but they have different semantics, different notions of conditional independence, etc.

Acknowledgment

This work has been supported in part by funds from grant GAČR 403/12/2175 to the first author, and from the Ronald G. Harper Distinguished Professorship at the University of Kansas to the second author. We are grateful to M. Studený for valuable discussions and comments. This paper is derived from a longer version in [7].

References

1. B. Ben-Yaghlane, P. Smets, and K. Mellouli. Belief function independence: II. The conditional case. *International Journal of Approximate Reasoning*, 31(1–2):31–75, 2002.
2. I. Couso, S. Moral, and P. Walley. Examples of independence for imprecise probabilities. In *Proceedings of the First International Symposium on Imprecise Probabilities and Their Applications (ISIPTA'99)*, 1999.
3. A. P. Dempster. Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics*, 38:325–339, 1967.
4. R. Jiroušek. Composition of probability measures on finite spaces. In D. Geiger and P. P. Shenoy, editors, *Uncertainty in Artificial Intelligence: Proceedings of the 13th Conference (UAI-97)*, pages 274–281, San Francisco, CA, 1997. Morgan Kaufmann.
5. R. Jiroušek. Foundations of compositional model theory. *International Journal of General Systems*, 40(6):623–678, 2011.
6. R. Jiroušek, J. Vejnarová, and M. Daniel. Compositional models of belief functions. In G. de Cooman, J. Vejnarová, and M. Zaffalon, editors, *Proceedings of the 5th Symposium on Imprecise Probabilities and Their Applications (ISIPTA-07)*, pages 243–252, Prague, Czech Republic, 2007. Charles University Press.
7. R. Jiroušek, and P. P. Shenoy. Compositional models in valuation-based systems. *Working Paper No. 325*, School of Business, University of Kansas, Lawrence, KS, 2011.
8. G. Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, 1976.
9. P. P. Shenoy. On Spohn's rule for revision of beliefs. *International Journal of Approximate Reasoning*, 5(2):149–181, 1991.
10. P. P. Shenoy. Conditional independence in valuation-based systems. *International Journal of Approximate Reasoning*, 10(3):203–234, 1994.
11. P. P. Shenoy. Binary join trees for computing marginals in the Shenoy-Shafer architecture. *International Journal of Approximate Reasoning*, 17(2–3):239–263, 1997.
12. P. P. Shenoy. No double counting semantics for conditional independence. In F. G. Cozman, R. Nau, and T. Seidenfeld, editors, *Proceedings of the Fourth International Symposium on Imprecise Probabilities and Their Applications (ISIPTA-05)*, pages 306–314. Society for Imprecise Probabilities and Their Applications, 2005.
13. R. P. Srivastava and K. Cogger. Beliefs on Individual Variables from a Single Source to Beliefs on the Joint Space under Dempster-Shafer Theory: An Algorithm. In J. Filipe, A. Fred, and B. Sharp, editors, *Proceedings of the First International Conference on Agents and Artificial Intelligence (ICAART 2009)*, INSTICC Press, Porto (Portugal), pages 191-197, 2009.
14. J. Vejnarová. Composition of possibility measures on finite spaces: Preliminary results. In B. Bouchon-Meunier and R. R. Yager, editors, *Proceedings of the 7th International Conference on Information Processing and Management of Uncertainty in Knowledge-based Systems (IPMU-98)*, pages 25–30, 1998.
15. P. Walley and T. L. Fine. Towards a frequentist theory of upper and lower probability. *Annals of Statistics*, 10:741–761, 1982.