On Two Operators of Composition in Dempster-Shafer Theory

Radim Jiroušek

Joint Research Laboratory of

Faculty of Management - Jindřichův Hradec — Inst. of Information Theory and Automation

University of Economics, Prague — Acad. of Sciences of the Czech republic

radim@utia.cas.cz

Composition in Probability Theory

Factorization Lemma.
For prob. distribution $\pi(X,Y,Z)$ there exist functions

$$\phi : \mathbb{X}\times\mathbb{Z} \longrightarrow \mathbb{R}^+,$$
$$\psi : \mathbb{Y}\times\mathbb{Z} \longrightarrow \mathbb{R}^+,$$

such that

$$\pi(\mathbf{a}) = \phi(\mathbf{a}^{\downarrow\{X,Z\}})\cdot\psi(\mathbf{a}^{\downarrow\{Y,Z\}})$$

for all $\mathbf{a}\in\mathbb{X}\times\mathbb{Y}\times\mathbb{Z}$ *iff* $X \perp\!\!\!\perp_\pi Y|Z$, and therefore also

$$\pi(X,Y,Z) = \pi(X,Z) \triangleright \pi(Y,Z).$$

Definition of Composition.
Let $\kappa(K)$ and $\lambda(L)$ be probability distributions.
If $\lambda^{\downarrow K\cap L} \gg \kappa^{\downarrow K\cap L}$ then

$$\kappa(K)\triangleright\lambda(L) = \frac{\kappa(K)\triangleright\lambda(L)}{\lambda(K\cap L)},$$

otherwise the composition is *undefined*.

Properties of Composition.

- Operator $\triangleright$ is neither commutative nor associative.
- $\kappa(K)\triangleright\lambda(L)$ is a prob. distribution of $K\cup L$.
- $\kappa(K)\triangleright\lambda(L)$ is an extension of κ.
- If κ and λ are consistent then $\kappa\triangleright\lambda$ is their joint extension.
- $\kappa(K)\triangleright\lambda(L)$ factorizes with respect to (K,L),
- and therefore also $K\setminus L \perp\!\!\!\perp_{\kappa\triangleright\lambda} L\setminus K|K\cap L$.
- If $X\in K$ and $X\notin L$ then $(\kappa\triangleright\lambda)^{-X} = \kappa^{-X}\triangleright\lambda$.

After defining the operator of composition in Shenoy's Valuation-Based Systems (VBS) [2] it appeared that there are (at least) two possibilities how to define an operator of composition in D-S theory:

Dempster operator of composition

Definition.
Consider two arbitrary bpas, μ_1 on $\mathbb{X}_K$ and μ_2 on $\mathbb{X}_L$ ($K\neq\emptyset\neq L$) and assume that $\mu_2^{\downarrow K\cap L} \gg \mu_1^{\downarrow K\cap L}$. Let θ_1 and θ_2 be their commonality functions, respectively. A *composition* $\theta_1 \overset{d}{\triangleright} \theta_2$ is defined for each nonempty $\mathbf{c}\subseteq\mathbb{X}_{K\cup L}$ by the following formula:

$$(\theta_1 \overset{d}{\triangleright} \theta_2)(\mathbf{c}) = \begin{cases} \alpha^{-1}\frac{\theta_1(\mathbf{c}^{\downarrow K})\cdot\theta_2(\mathbf{c}^{\downarrow L})}{\theta_2^{\downarrow K\cap L}(\mathbf{c}^{\downarrow K\cap L})} & \text{if } \theta_2^{\downarrow K\cap L}(\mathbf{c}^{\downarrow K\cap L})>0,\\ 0 & \text{otherwise,}\end{cases}$$

where α is a normalization constant defined as

$$\alpha = \sum_{\mathbf{d}\in 2^{\mathbb{X}_{K\cup L}}:\theta_2^{\downarrow K\cap L}(\mathbf{d}^{\downarrow K\cap L})>0} (-1)^{|\mathbf{d}|+1}\frac{\theta_1(\mathbf{d}^{\downarrow K})\cdot\theta_2(\mathbf{d}^{\downarrow L})}{\theta_2^{\downarrow K\cap L}(\mathbf{d}^{\downarrow K\cap L})}.$$

Factorizing operator of composition

Definition.
Consider two normal bpas, μ_1 on $\mathbb{X}_K$ and μ_2 on $\mathbb{X}_L$ ($K\neq\emptyset\neq L$). A *composition* $\mu_1 \overset{f}{\triangleright} \mu_2$ is defined for each nonempty $\mathbf{c}\subseteq\mathbb{X}_{K\cup L}$ by one of the following expressions:

(i) if $\mu_2^{\downarrow K\cap L}(\mathbf{c}^{\downarrow K\cap L})>0$ and $\mathbf{c}=\mathbf{c}^{\downarrow K}\bowtie\mathbf{c}^{\downarrow L}$ then

$$(\mu_1 \overset{f}{\triangleright} \mu_2)(\mathbf{c}) = \frac{\mu_1(\mathbf{c}^{\downarrow K})\cdot\mu_2(\mathbf{c}^{\downarrow L})}{\mu_2^{\downarrow K\cap L}(\mathbf{c}^{\downarrow K\cap L})};$$

(ii) if $\mu_2^{\downarrow K\cap L}(\mathbf{c}^{\downarrow K\cap L})=0$ and $\mathbf{c}=\mathbf{c}^{\downarrow K}\times\mathbb{X}_{L\setminus K}$ then

$$(\mu_1 \overset{f}{\triangleright} \mu_2)(\mathbf{c}) = m_1(\mathbf{c}^{\downarrow K});$$

(iii) in all other cases, $(\mu_1 \overset{f}{\triangleright} \mu_2)(\mathbf{c})=0$.

Properties of the Dempster operator

- Meets all the properties of composition (can be used for multidimensional model representation, makes local computations in decomposable models possible).
- Computationally **extremely complex**.
- Can be used in the process of Iterative Proportional Fitting but **Csiszár's convergence theorem [1] does not hold**.
- **Defines conditionals:** for bpa μ on $\mathbb{X}_M$ and $X_j, X_k\in M$,

$$\mu(X_k|X_j=\mathbf{a}) = \left(\nu_{X_j=\mathbf{a}} \overset{d}{\triangleright} \mu\right)^{\downarrow X_k},$$

  where $\nu_{X_j=\mathbf{a}}$ is a one-dimensional bpa on $\mathbb{X}_j$ having just one focal element $\{\mathbf{a}\}\subset\mathbb{X}_j$, for which $\nu_{X_j=\mathbf{a}}(\{\mathbf{a}\})=1$.
- **Conditional Independence Lemma.**
  For commonality function $\theta(X,Y,Z)$ the conditional independence relation $X \perp\!\!\!\perp_\theta Y|Z$ holds true *iff*

$$\theta(X,Y,Z) = \theta(X,Z) \overset{d}{\triangleright} \theta(Y,Z).$$

Properties of the factorizintg operator

- Meets all the properties of composition (can be used for multidimensional model representation, makes local computations in decomposable models possible).
- Computationally **much simpler** than Dempster operator.
- Can be used in the process of Iterative Proportional Fitting; **Csiszár's convergence theorem [1] holds true.**
- **Cannot be used to express conditionals.**
- **Factorization Lemma.**
  For bpa $\mu(X,Y,Z)$ there exist functions

$$\phi : 2^{\mathbb{X}\times\mathbb{Z}} \longrightarrow \mathbb{R}^+,\qquad \psi : 2^{\mathbb{Y}\times\mathbb{Z}} \longrightarrow \mathbb{R}^+,$$

  such that

$$\mu(\mathbf{a}) = \begin{cases}\phi(\mathbf{a}^{\downarrow\{X,Z\}})\cdot\psi(\mathbf{a}^{\downarrow\{Y,Z\}}) & \text{if } \mathbf{a}=\mathbf{a}^{\downarrow\{X,Z\}}\bowtie\mathbf{a}^{\downarrow\{Y,Z\}}\\ 0 & \text{otherwise}\end{cases}$$

  iff $\mu(X,Y,Z) = \mu(X,Z) \overset{f}{\triangleright} \mu(Y,Z)$.

Main Open Problem - Conjecture

Suppose μ_1, μ_2 and μ_3 are bpas on $\mathbb{X}_K$, $\mathbb{X}_L$, and $\mathbb{X}_M$, respectively. If $L\supset(K\cap M)$ then, $(\mu_1 \overset{d}{\triangleright} \mu_2) \overset{f}{\triangleright} \mu_3 = \mu_1 \overset{d}{\triangleright} (\mu_2 \overset{f}{\triangleright} \mu_3)$.

References

[1] I. Csiszár. I-divergence geometry of probability distributions and minimization problems. *Ann. Probab.* 3, pp 146–158, 1975.

[2] R. Jiroušek and P. P. Shenoy. Compositional models in valuation-based systems. *Int. J. Approx. Reasoning*, 55, 1, pp. 277–293, 2014.