Almost Bayesian Assignments and Conditional Independence

(a contribution to Dempster-Shafer theory of evidence)

Radim Jiroušek

Joint Research Laboratory of
Faculty of Management - Jindřichův Hradec, University of Economics, Prague
Inst. of Information Theory and Automation, Acad. of Sciences of the Czech republic

radim@utia.cas.cz

Set notation

- *Frame of discernment* $\mathbf{X}_N = \mathbf{X}_1 \times \mathbf{X}_2 \times \ldots \times \mathbf{X}_n$;

For $K, L, \ldots \subseteq N$ $\quad \mathbf{X}_K = \times_{i\in K}\mathbf{X}_i$.

- By a *join* of two sets $A \subseteq \mathbf{X}_K$ and $B \subseteq \mathbf{X}_L$ we will denote

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

Notice that if: K and L are disjoint $\implies A \otimes B = A \times B$.

$K = L \implies A \otimes B = A \cap B$.

Operator of composition

Principle For basic assignments m_1 on $\mathbf{X}_K$ and m_2 on $\mathbf{X}_L$ their *composition* $m_1 \triangleright m_2$ is a basic assignment on $\mathbf{X}_{K\cup L}$ defined in the way that the mass $m_1(C)$ is *disseminated* into subsets of $C \times \mathbf{X}_{L\setminus K}$.

Definition 1 For 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 $\boxed{m_2^{\downarrow K\cap L}(C^{\downarrow K\cap L}) > 0 \text{ and } C = C^{\downarrow K} \otimes C^{\downarrow L}}$ then

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

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

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

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

Theorem 1 Let $K, L \subseteq N$. For basic assignments m_1, m_2 defined on $\mathbf{X}_K$, $\mathbf{X}_L$, respectively

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

Remarks

- Highlighted are the properties, which do not hold for the Dempster's rule of combination.
- Generally, the operation $\triangleright$ is neither commutative nor associative.
- The operation of composition is *idempotent*: $m \triangleright m = m$.

Compositional models

Goal Having a number of low-dimensional basic assignments $m_1, m_2, \ldots, m_\ell$ defined on $\mathbf{X}_{K_1}, \mathbf{X}_{K_2}, \ldots, \mathbf{X}_{K_\ell}$, respectively, the goal is to obtain a multi-dimensional basic assignment.

Solution Take a (suitable) ordering $m_{j_1}, m_{j_2}, \ldots, m_{j_\ell}$, and consider

$$m_{j_1} \triangleright m_{j_2} \triangleright m_{j_3} \triangleright \ldots \triangleright m_{j_\ell} = (\ldots((m_{j_1} \triangleright m_{j_2}) \triangleright m_{j_3}) \triangleright \ldots \triangleright m_{j_{\ell-1}}) \triangleright m_{j_\ell}$$

Definition 2 Compositional model $m_1 \triangleright m_2 \triangleright \ldots \triangleright m_\ell$ is called *perfect sequence model* if for all $j = 2, \ldots, \ell$

$$\boxed{m_1 \triangleright \ldots \triangleright m_{j-1} \triangleright m_j = m_j \triangleright (m_1 \triangleright \ldots \triangleright m_{j-1}).}$$

Theorem 2 Compositional model $m_1 \triangleright m_2 \triangleright \ldots \triangleright m_\ell$ is a perfect sequence model *iff* each m_j is marginal to $m_1 \triangleright m_2 \triangleright \ldots \triangleright m_\ell$.

Almost Bayesian basic assignments

Definition 3 Basic assignment m is said to be *almost Bayesian* if it is

- *cylindrical* - all its focal elements C are point-cylinders ($C = C^{\downarrow L} \times \mathbf{X}_{K\setminus L}$ for $|C^{\downarrow L}| \leq 1$); and
- *sparse* (*quasi-Bayesian*)- all its focal elements are pairwise disjoint.

Theorem 3 If both basic assignments m_1 and m_2 are cylindrical (sparse) then also their composition $m_1 \triangleright m_2$ is cylindrical (sparse).

Conclusion Since each Bayesian basic assignment (i.e. basic assignment whose focal elements are singletons) is both cylindrical and sparse it is obvious that any compositional model assembled from Bayesian basic assignments is almost Bayesian.

Remark It is known that a **perfect** sequence model assembled from Bayesian basic assignments is Bayesian.

Conditional Independence

Motivation In probability theory conditional independence is tightly connected with *factorization*:

$$X \perp\!\!\!\perp Y | Z \, [\pi] \iff \pi(X,Y,Z) = \frac{\pi(X,Z)\cdot\pi(Y,Z)}{\pi(Z)} = \pi(X,Z) \triangleright \pi(Y,Z).$$

Definition 4 For a basic assignment m on $\mathbf{X}_N$ and disjoint $K, L, M \subset N$ we say that $X_K \perp\!\!\!\perp X_L | X_M \, [m]$ if

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

Comparison with conditional non-interactivity of Ben Yaghlane *et al.* (2002)

- Both conditional independence and conditional non-interactivity meets *semi-graphoid axioms.*
- **Unconditional** independence and **unconditional** non-interactivity coincide.
- For **Bayesian** basic assignments both conditional independence and conditional non-interactivity coincide with probabilistic conditional independence.
- If for basic assignment m variables X_K and X_L are conditionally independent (non-interactive) given variables X_M then for each $C \subset \mathbf{X}_{K\cup L\cup M}$

$$C \neq C^{\downarrow K\cup M} \otimes C^{\downarrow L\cup M} \implies m(C) = 0.$$

- **Conditional independence is *consistent with marginalization*: having two consistent basic assignments m_1 on $\mathbf{X}_K$ and m_2 on $\mathbf{X}_L$ one can always find their common extension on $\mathbf{X}_{K\cup L}$, for which $X_{K\setminus L} \perp\!\!\!\perp X_{L\setminus K} | X_{K\cap L}$. (It is $m_1 \triangleright m_2$.)**

Comparison of the operator of composition and Dempster's rule of combination

- Operator of composition (in contrast to Dempster's rule of combination) is neither commutative nor associative.
- Dempster's rule of combination does not meet properties 2. and 4. of Theorem 1.

Jiroušek, R. and Vejnarová, J. (2009). There are Combinations and Compositions in Dempster-Shafer Theory of Evidence. *To appear in:* Proceedings of Workshop on Uncertainty Processing, WUPES'09.

Summary Operator of composition and Dempster's rule of combination coincide only in very special situations. For example, $m_1 \oplus m_2 = m_1 \triangleright m_2$

- when the combined basic assignments m_1 and m_2 are defined on disjoint frames of discernment;
- if they are defined on the same frame of discernment and all focal elements of m_1 are contained in each focal element of m_2;
- when all the focal elements of both m_1 and m_2 project to the same subset of the intersecting frame of discernment.