Almost Bayesian Assignments and Conditional Independence

Radim Jiroušek

Institute of Information Theory and Automation
Academy of Sciences of the Czech Republic
and
Faculty of Management
University of Economics
Jindřichův Hradec

Outline of the Lecture

History

1986 Jiroušek Radim, Perez Albert: Graph-aided Knowledge Integration in Expert System INES. Proceedings IPMU'86.

1997 Jiroušek Radim: Composition of probability measures on finite spaces. Proceedings UAI'97.

1998 Vejnarová Jirina: Composition of possibility measures on finite spaces: Preliminary results. Proceedings IPMU'98.

2007 Jiroušek Radim, Vejnarová Jirina, Daniel Milan: Compositional models for belief functions. Proceedings ISIPTA'07.

Probabilistic operator of composition

For $\kappa_1(x_K)$ and $\kappa_2(x_L)$ defined on $\mathbf{X}_K$ and $\mathbf{X}_L$, respectively, such that $\kappa_1^{\downarrow K\cap L} \ll \kappa_2^{\downarrow K\cap L}$, which means that

$$\forall x \in \mathbf{X}_{K\cap L} \quad (\kappa_2^{\downarrow K\cap L}(x) = 0 \Longrightarrow \kappa_1^{\downarrow K\cap L}(x) = 0);$$

their composition is defined for all $x \in \mathbf{X}_{K\cup L}$

$$(\kappa_1 \triangleright \kappa_2)(x) = \frac{\kappa_1(x^{\downarrow K})\kappa_2(x^{\downarrow L})}{\kappa_2^{\downarrow K\cap L}(x^{\downarrow K\cap L})} = \kappa_1(x^{\downarrow K})\kappa_2(x^{\downarrow L\setminus K}|x^{\downarrow L\cap K}).$$

Basic properties of the operator of composition

- $\kappa_1(x_K) \triangleright \kappa_2(x_L) = (\kappa_1 \triangleright \kappa_2)(x_{K \cup L})$;
- $(\kappa_1(x_K) \triangleright \kappa_2(x_L))^{\downarrow K} = \kappa_1(x_K)$;
- operator is neither commutative nor associative;
- $\kappa_1 \triangleright \kappa_2 = \kappa_2 \triangleright \kappa_1 \iff \kappa_1^{\downarrow K \cap L} = \kappa_2^{\downarrow K \cap L}$;
- $X_{K \setminus L} \perp\!\!\!\perp X_{L \setminus K} | X_{K \cap L} \; [\kappa_1 \triangleright \kappa_2]$;
- $\ldots$;
- $X_I \perp\!\!\!\perp X_J | X_K[\kappa] \iff \kappa(x^{\downarrow I \cup J \cup K}) = \kappa(x^{\downarrow I \cup K}) \triangleright \kappa(x^{\downarrow J \cup K})$.

Basic properties of the operator of composition

- $\kappa_1(x_K) \triangleright \kappa_2(x_L) = (\kappa_1 \triangleright \kappa_2)(x_{K \cup L})$;
- $(\kappa_1(x_K) \triangleright \kappa_2(x_L))^{\downarrow K} = \kappa_1(x_K)$;
- operator is neither commutative nor associative;
- $\kappa_1 \triangleright \kappa_2 = \kappa_2 \triangleright \kappa_1 \iff \kappa_1^{\downarrow K \cap L} = \kappa_2^{\downarrow K \cap L}$;
- $X_{K \setminus L} \perp\!\!\!\perp X_{L \setminus K} | X_{K \cap L} \; [\kappa_1 \triangleright \kappa_2]$;
- $\ldots$;
- $X_I \perp\!\!\!\perp X_J | X_K[\kappa] \iff \kappa(x^{\downarrow I \cup J \cup K}) = \kappa(x^{\downarrow I \cup K}) \triangleright \kappa(x^{\downarrow J \cup K})$.

Basic properties of the operator of composition

- $\kappa_1(x_K) \triangleright \kappa_2(x_L) = (\kappa_1 \triangleright \kappa_2)(x_{K\cup L})$;
- $(\kappa_1(x_K) \triangleright \kappa_2(x_L))^{\downarrow K} = \kappa_1(x_K)$;
- operator is neither commutative nor associative;
- $\kappa_1 \triangleright \kappa_2 = \kappa_2 \triangleright \kappa_1 \iff \kappa_1^{\downarrow K\cap L} = \kappa_2^{\downarrow K\cap L}$;
- $X_{K\setminus L} \perp\!\!\!\perp X_{L\setminus K} | X_{K\cap L}\ [\kappa_1 \triangleright \kappa_2]$;
- $\ldots$;
- $X_I \perp\!\!\!\perp X_J | X_K [\kappa] \iff \kappa(x^{\downarrow I\cup J\cup K}) = \kappa(x^{\downarrow I\cup K}) \triangleright \kappa(x^{\downarrow J\cup K})$.

Basic properties of the operator of composition

- $\kappa_1(x_K) \triangleright \kappa_2(x_L) = (\kappa_1 \triangleright \kappa_2)(x_{K\cup L})$;
- $(\kappa_1(x_K) \triangleright \kappa_2(x_L))^{\downarrow K} = \kappa_1(x_K)$;
- operator is neither commutative nor associative;
- $\kappa_1 \triangleright \kappa_2 = \kappa_2 \triangleright \kappa_1 \iff \kappa_1^{\downarrow K\cap L} = \kappa_2^{\downarrow K\cap L}$;
- $X_{K\setminus L} \perp\!\!\!\perp X_{L\setminus K} | X_{K\cap L}\ [\kappa_1 \triangleright \kappa_2]$;
- ...;
- $X_I \perp\!\!\!\perp X_J | X_K[\kappa] \iff \kappa(x^{\downarrow I\cup J\cup K}) = \kappa(x^{\downarrow I\cup K}) \triangleright \kappa(x^{\downarrow J\cup K})$.

Generating sequences

$$\kappa_1(x_{K_1}), \kappa_2(x_{K_2}), \ldots, \kappa_n(x_{K_n})$$

Definition

Compositional Model:

$$\kappa_1(x_{K_1}) \triangleright \kappa_2(x_{K_2}) \triangleright \ldots \triangleright \kappa_n(x_{K_n})$$
$$= \left(\ldots\left(\kappa_1(x_{K_1}) \triangleright \kappa_2(x_{K_2})\right) \triangleright \ldots \triangleright \kappa_n(x_{K_n})\right)$$

Generating sequences

$$\kappa_1(x_{K_1}), \kappa_2(x_{K_2}), \ldots, \kappa_n(x_{K_n})$$

Definition

Compositional Model:

$$\kappa_1(x_{K_1}) \triangleright \kappa_2(x_{K_2}) \triangleright \ldots \triangleright \kappa_n(x_{K_n})$$

$$= \left(\ldots\left(\kappa_1(x_{K_1}) \triangleright \kappa_2(x_{K_2})\right) \triangleright \ldots \triangleright \kappa_n(x_{K_n})\right)$$

Generating sequences

$$\kappa_1(x_{K_1}), \kappa_2(x_{K_2}), \ldots, \kappa_n(x_{K_n})$$

Definition

Compositional Model:

$$\kappa_1(x_{K_1}) \triangleright \kappa_2(x_{K_2}) \triangleright \ldots \triangleright \kappa_n(x_{K_n})$$

$$= \left(\ldots\left(\kappa_1(x_{K_1}) \triangleright \kappa_2(x_{K_2})\right) \triangleright \ldots \triangleright \kappa_n(x_{K_n})\right)$$

Perfect sequences

Definition

Generating sequence $\kappa_1(x_{K_1}), \kappa_2(x_{K_2}), \ldots, \kappa_n(x_{K_n})$ is perfect if

$$\kappa_1 \triangleright \kappa_2 = \kappa_2 \triangleright \kappa_1,$$

$$(\kappa_1 \triangleright \kappa_2) \triangleright \kappa_3 = \kappa_3 \triangleright (\kappa_1 \triangleright \kappa_2),$$

$$(\kappa_1 \triangleright \kappa_2 \triangleright \kappa_3) \triangleright \kappa_4 = \kappa_4 \triangleright (\kappa_1 \triangleright \kappa_2 \triangleright \kappa_3),$$

$$\cdots$$

$$(\kappa_1 \triangleright \ldots \triangleright \kappa_{n-1}) \triangleright \kappa_n = \kappa_n \triangleright (\kappa_1 \triangleright \ldots \triangleright \kappa_{n-1}).$$

Perfect sequences

Characterization Theorem:

Generating sequence $\kappa_1, \kappa_2, \ldots, \kappa_n$ is perfect iff all κ_i are marginal distributions of $\kappa_1 \triangleright \kappa_2 \triangleright \ldots \triangleright \kappa_n$.

Comparison with Bayesian networks

Both compositional models and Bayesian networks represent the same class of distributions

Pros

- It does not need help of Graph Theory
- Perfect sequence models are computationally more efficient

Cons

- Compositional models are not always defined
- It is more difficult to construct perfect sequence models

Operator of composition was defined also in Possibility theory (1998) and D-S theory of evidence (2007).

Comparison with Bayesian networks

Both compositional models and Bayesian networks represent the same class of distributions

Pros

- It does not need help of Graph Theory
- Perfect sequence models are computationally more efficient

Cons

- Compositional models are not always defined
- It is more difficult to construct perfect sequence models

Operator of composition was defined also in Possibility theory (1998) and D-S theory of evidence (2007).

Comparison with Bayesian networks

Both compositional models and Bayesian networks represent the same class of distributions

Pros

- It does not need help of Graph Theory
- Perfect sequence models are computationally more efficient

Cons

- Compositional models are not always defined
- It is more difficult to construct perfect sequence models

Operator of composition was defined also in Possibility theory (1998) and D-S theory of evidence (2007).

Comparison with Bayesian networks

Both compositional models and Bayesian networks represent the same class of distributions

Pros

- It does not need help of Graph Theory
- Perfect sequence models are computationally more efficient

Cons

- Compositional models are not always defined
- It is more difficult to construct perfect sequence models

Operator of composition was defined also in Possibility theory (1998) and D-S theory of evidence (2007).

Set notation

Let $K \subset L \subseteq N$ and $x \in \mathbf{X}_L$. $x^{\downarrow K}$ denotes a *projection* of x into $\mathbf{X}_K$.

Analogously, for and $A \subset \mathbf{X}_L$, $A^{\downarrow K}$ denotes a *projection* of A into $\mathbf{X}_K$:

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

Important: we do not exclude $K = \emptyset$. In this case $A^{\downarrow \emptyset} = \emptyset$.

A *join* of two sets $A \subseteq \mathbf{X}_K$ and $B \subseteq \mathbf{X}_M$ is the set

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

Operator of composition

For basic assignments m_1 on $\mathbf{X}_K$ and m_2 on $\mathbf{X}_L$ a *composition* $m_1 \triangleright m_2$ is defined for all $C \subseteq \mathbf{X}_{K \cup L}$ by one of the following expressions:

[a] if $m_2^{\downarrow K\cap L}(C^{\downarrow K\cap L}) > 0$ and $C = C^{\downarrow K} \otimes C^{\downarrow L}$ 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 $m_2^{\downarrow K\cap L}(C^{\downarrow K\cap L}) = 0$ 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 $(m_1 \triangleright m_2)(C) = 0.$

Basic properties

For ba's m_1 on $\mathbf{X}_K$ and m_2 on $\mathbf{X}_L$:

- $m_1 \triangleright m_2$ is ba on $\mathbf{X}_{K \cup L}$;
- m_1 is marginal of $m_1 \triangleright m_2$;
- operator is neither commutative nor associative;
- ....;
- for all $A \subseteq \mathbf{X}^{\downarrow K \cup L}, A \neq A^{\downarrow K} \otimes A^{\downarrow L} \implies (m_1 \triangleright m_2)(A) = 0$;

E.g. for binary case $|\mathbf{X}_{\{1,2,3\}}| = 2^8 - 1 = 255$

$$|\{A : A = A^{\downarrow \{1,2\}} \otimes A^{\downarrow \{2,3\}}\}| = 99.$$

Basic properties

For ba's m_1 on $\mathbf{X}_K$ and m_2 on $\mathbf{X}_L$:

- $m_1 \triangleright m_2$ is ba on $\mathbf{X}_{K \cup L}$;
- m_1 is marginal of $m_1 \triangleright m_2$;
- operator is neither commutative nor associative;
- ....;
- for all $A \subseteq \mathbf{X}^{\downarrow K \cup L}, A \neq A^{\downarrow K} \otimes A^{\downarrow L} \implies (m_1 \triangleright m_2)(A) = 0$;

E.g. for binary case $|\mathbf{X}_{\{1,2,3\}}| = 2^8 - 1 = 255$

$$|\{A : A = A^{\downarrow \{1,2\}} \otimes A^{\downarrow \{2,3\}}\}| = 99.$$

Almost Bayesian basic assignments

Definition

Basic assignment 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.

Assertion

Any compositional model assembled from Bayesian basic assignments is almost Bayesian.

Almost Bayesian basic assignments

Definition

Basic assignment 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.

Assertion

Any compositional model assembled from Bayesian basic assignments is almost Bayesian.

Conditional noninteractivity for belief functions

There are several definitions of this notion:
The most frequent is that used by M. Studený, P. Shenoy, Ben Yaghlane et al.:

Definition

Variables X_I and X_J are *conditionally non-interactive* given variables X_K ($X_I \perp\!\!\!\perp_{[m]} X_J | X_K$) if for all $A \subseteq \mathbf{X}_N$

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

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

Conditional independence for belief functions

In this case, however:

$$X_I \perp\!\!\!\perp_{[m]} X_J | X_K \quad \not\Longleftrightarrow \quad m^{\downarrow I\cup J\cup K} = m^{\downarrow I\cup K} \triangleright m^{\downarrow J\cup K}$$

If we substitute conditional irrelevance by factorization we get conditional independence relation with the following properties:

- its restriction to Bayesian basic assignment corresponds to probabilistic conditional independence relation;
- it meets all the semigraphoid axioms (symmetry and Block independence lemma);
- it does not suffer from the conditional irrelevance imperfectness: it is consistent with marginalization.

Conditional independence for belief functions

In this case, however:

$$X_I \perp\!\!\!\perp_{[m]} X_J | X_K \quad \not\Leftrightarrow \quad m^{\downarrow I \cup J \cup K} = m^{\downarrow I \cup K} \triangleright m^{\downarrow J \cup K}$$

If we substitute conditional irrelevance by factorization we get conditional independence relation with the following properties:

- its restriction to Bayesian basic assignment corresponds to probabilistic conditional independence relation;
- it meets all the semigraphoid axioms (symmetry and Block independence lemma);
- it does not suffer from the conditional irrelevance imperfectness: it is consistent with marginalization.

Conditional independence for belief functions

In this case, however:

$$X_I \perp\!\!\!\perp_{[m]} X_J | X_K \quad \not\Longleftrightarrow \quad m^{\downarrow I\cup J\cup K} = m^{\downarrow I\cup K} \triangleright m^{\downarrow J\cup K}$$

If we substitute conditional irrelevance by factorization we get conditional independence relation with the following properties:

- its restriction to Bayesian basic assignment corresponds to probabilistic conditional independence relation;
- it meets all the semigraphoid axioms (symmetry and Block independence lemma);
- it does not suffer from the conditional irrelevance imperfectness: it is consistent with marginalization.

Conditional independence for belief functions

In this case, however:

$$X_I \perp\!\!\!\perp_{[m]} X_J | X_K \quad \not\Leftrightarrow \quad m^{\downarrow I \cup J \cup K} = m^{\downarrow I \cup K} \triangleright m^{\downarrow J \cup K}$$

If we substitute conditional irrelevance by factorization we get conditional independence relation with the following properties:

- its restriction to Bayesian basic assignment corresponds to probabilistic conditional independence relation;
- it meets all the semigraphoid axioms (symmetry and Block independence lemma);
- it does not suffer from the conditional irrelevance imperfectness: it is consistent with marginalization.

Consistency with marginalization

Having two projective ba's m_1 on $\mathbf{X}_K$ and m_2 on $\mathbf{X}_M$, (i.e. $m_1^{\downarrow K\cap M} = m_2^{\downarrow K\cap M}$) does there exist a ba m such that:

- m_1 and m_2 are marginal assignments of m;
- $X_{K\setminus M} \perp\!\!\!\perp_{[m]} X_{M\setminus K} | X_{K\cap M}$?

The solution is simple:

$$m = m_1 \triangleright m_2.$$

Consistency with marginalization

Having two projective ba's m_1 on $\mathbf{X}_K$ and m_2 on $\mathbf{X}_M$,
(i.e. $m_1^{\downarrow K\cap M} = m_2^{\downarrow K\cap M}$) does there exist a ba m such that:

- m_1 and m_2 are marginal assignments of m;
- $X_{K\setminus M} \perp\!\!\!\perp_{[m]} X_{M\setminus K} | X_{K\cap M}$?

The solution is simple:

$$m = m_1 \triangleright m_2.$$

Consistency with marginalization

Having two projective ba's m_1 on $\mathbf{X}_K$ and m_2 on $\mathbf{X}_M$, (i.e. $m_1^{\downarrow K\cap M} = m_2^{\downarrow K\cap M}$) does there exist a ba m such that:

- m_1 and m_2 are marginal assignments of m;
- $X_{K\setminus M} \perp\!\!\!\perp_{[m]} X_{M\setminus K} | X_{K\cap M}$?

The solution is simple:

$$m = m_1 \triangleright m_2.$$

THANK YOU FOR YOUR ATTENTION