On Two Approaches to Evidential Network Construction

Jiřina Vejnarová*

Institute of Information Theory and
Automation of the ASCR
&
University of Economics, Prague
vejnar@utia.cas.cz

Abstract: We compare two approaches to the construction of belief networks in the framework of evidence theory. We show that belief networks reconstructed from multidimensional models based on recently introduced operator of composition of basic assignments have similar structural properties to Bayesian networks. On the other hand, it is not valid for previously proposed directed evidential networks by Ben Yaghlane at al. The difference between the structural properties of models based on these two approaches is illustrated by a simple example.

Keywords: belief network, conditional independence, conjunctive combination rule

1 Introduction

Bayesian networks are at present the most popular representative of so-called graphical Markov models. Therefore it is not surprising that some attempts to construct an analogy of Bayesian networks have also been made in other frameworks as e.g. in possibility theory [4] or evidence theory [3].

Since the method from [3] does not seem to us to be satisfactory, as graphical tools well-known from Bayesian networks are used in different sense, therefore in [9] we brought an alternative to it. Our approach is based on recently introduced operator of composition for basic assignments [6]. The evidential network is reconstructed from the resulting compositional model. We concentrate ourselves to structural properties of the network, the problem of definition of conditional beliefs has not been solved yet. The main contribution of this paper is an illustrative example showing the differences between these two approaches

The paper is organized as follows. After a brief summary of basic notions from evidence theory (Section 2), in Section 3 we recall the definition of the operator of composition, perfect sequences of basic assignments and an algorithm for transformation of a perfect sequence into an evidential network. Section 4 is devoted to a simple example comparing our approach with the previous one [3].

2 Basic notions

In this section we will briefly recall basic concepts from evidence theory concerning sets, set functions and (conditional) independence.

*The work of the author was supported by the grant GA ČR 201/09/1891, by the grant GA AV ČR A100750603 and by the grant MŠMT 2C06019.

2.1 Set projections and joins

For an index set $N = \{1, 2, \ldots, n\}$ let $\{X_i\}_{i \in N}$ be a system of variables, each X_i having its values in a finite set $\mathbf{X}_i$. In this paper we will deal with *multidimensional frame of discernment* $\mathbf{X}_N = \mathbf{X}_1 \times \mathbf{X}_2 \times \ldots \times \mathbf{X}_n$, and its *subframes* (for $K \subseteq N$) $\mathbf{X}_K = \times_{i \in K} \mathbf{X}_i$. When dealing with groups of variables on these subframes, X_K will denote a group of variables $\{X_i\}_{i \in K}$ throughout the paper.

A *projection* $x^{\downarrow K}$ of $x = (x_1, x_2, \ldots, x_n) \in \mathbf{X}_N$ into $\mathbf{X}_K$ is for $K = \{i_1, i_2, \ldots, i_k\}$

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

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

$$A^{\downarrow M} = \{y \in \mathbf{X}_M \mid \exists x \in A : y = x^{\downarrow M}\}.$$

In addition to the projection, in this text we will also need an opposite operation, which will be called a join. By a *join** of two sets $A \subseteq \mathbf{X}_K$ and $B \subseteq \mathbf{X}_L$ $(K, L \subseteq N)$ we will understand a set

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

Let us note that for any $C \subseteq \mathbf{X}_{K \cup L}$ naturally $C \subseteq C^{\downarrow K} \bowtie C^{\downarrow L}$, but generally $C \neq C^{\downarrow K} \bowtie C^{\downarrow L}$.

2.2 Set functions

In evidence theory (or Dempster-Shafer theory) two measures are used to model the uncertainty: belief and plausibility measures (the latter one will not be used in this paper). Both of them can be defined with the help of another set function called a *basic (probability* or *belief) assignment* m on $\mathbf{X}_N$, i.e.,

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

where $\mathcal{P}(\mathbf{X}_N)$ is power set of $\mathbf{X}_N$ and $\sum_{A \subseteq \mathbf{X}_N} m(A) = 1$. Furthermore, we assume (in contrary to [2, 3]) that $m(\emptyset) = 0$.

Belief measure is defined for any $A \subseteq \mathbf{X}_N$ by the equality

$$Bel(A) \quad = \quad \sum_{B \subseteq A} m(B). \qquad (1)$$

For a basic assignment m on $\mathbf{X}_K$ and $M \subset K$, a *marginal basic assignment* of m on $\mathbf{X}_M$ is defined (for each $A \subseteq \mathbf{X}_M$):

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

Having two basic assignments m_1 and m_2 on $\mathbf{X}_K$ and $\mathbf{X}_L$, respectively $(K, L \subseteq N)$, we say that these assignments are *projective* if

$$m_1^{\downarrow K \cap L} = m_2^{\downarrow K \cap L},$$

which occurs if and only if there exists a basic assignment m on $\mathbf{X}_{K \cup L}$ such that both m_1 and m_2 are marginal assignments of m. Let us note that according to the convention $m^{\downarrow \emptyset} \equiv 1$ for arbitrary basic assignment m, m_1 and m_2 are projective whenever $K \cap L = \emptyset$.

*This term and notation are taken from the theory of relational databases [1].

2.3 Independence

When constructing graphical models in any framework, (conditional) independence concept plays an important role. In evidence theory the most common notion of independence is that of random set independence [5]: Let m be a basic assignment on $\mathbf{X}_N$ and $K, L \subset N$ be disjoint. We say that groups of variables X_K and X_L are *independent with respect to basic assignment* m (in notation $K \perp\!\!\!\perp L\ [m]$) if

$$m^{\downarrow K \cup L}(A) = m^{\downarrow K}(A^{\downarrow K}) \cdot m^{\downarrow L}(A^{\downarrow L})$$

for all $A \subseteq \mathbf{X}_{K \cup L}$ for which $A = A^{\downarrow K} \times A^{\downarrow L}$, and $m(A) = 0$ otherwise.

This notion can be generalized in various ways [7, 2, 8]; the concept of conditional non-interactivity $X_K \perp_m X_L | X_M$ from [2], based on conjunction combination rule, is used for construction of directed evidential networks in [3]. In this paper we will use the concept introduced in [8, 6], as we consider it more suitable (the arguments can be found in [8]).

Definition 1 Let m be a basic assignment on $\mathbf{X}_N$ and $K, L, M \subset N$ be disjoint, $K \neq \emptyset \neq L$. We say that groups of variables X_K and X_L are *conditionally independent given* X_M *with respect to* m (and denote it by $K \perp\!\!\!\perp L | M\ [m]$), if the equality

$$m^{\downarrow K \cup L \cup M}(A) \cdot m^{\downarrow M}(A^{\downarrow M}) = m^{\downarrow K \cup M}(A^{\downarrow K \cup M}) \cdot m^{\downarrow L \cup M}(A^{\downarrow L \cup M}) \tag{2}$$

holds for any $A \subseteq \mathbf{X}_{K \cup L \cup M}$ such that $A = A^{\downarrow K \cup M} \bowtie A^{\downarrow L \cup M}$, and $m(A) = 0$ otherwise.

It has been proven in [8] that this conditional independence concept satisfies so-called semi-graphoid properties taken as reasonable to be valid for any conditional independence concept.

3 Belief network generated by a perfect sequence

3.1 Operator of composition

Operator of composition of basic assignments was introduced in [6] in the following way.

Definition 2 *For two arbitrary 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* $C = C^{\downarrow K} \bowtie C^{\downarrow L}$ *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*

$$(m_1 \triangleright m_2)(C) = 0.$$

A lot of properties possessed by the operator of composition can be found in [6], nevertheless here we will confine ourselves to the following theorem (proven in [6]) expressing the relationship between conditional independence (from Definition 1) and operator of composition.

Theorem 3 *Let* m *be a joint basic assignment on* $\mathbf{X}_M$, $K, L \subseteq M$. *Then* $(K \setminus L) \perp\!\!\!\perp (L \setminus K) | (K \cap L)\ [m]$ *if and only if*

$$m^{\downarrow K \cup L}(A) = (m^{\downarrow K} \triangleright m^{\downarrow L})(A)$$

for any $A \subseteq \mathbf{X}_{K \cup L}$.

3.2 Perfect sequences

Now, let us consider a system of low-dimensional 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. Composing them together by multiple application of the operator of composition, one gets multidimensional basic assignments on $\mathbf{X}_{K_1 \cup K_2 \cup \ldots \cup K_n}$.

To avoid using too many parentheses let us make the convention, that the operator of composition is performed successively from left to right:

$$m_1 \triangleright m_2 \triangleright \ldots \triangleright m_n = (\ldots((m_1 \triangleright m_2) \triangleright m_3) \triangleright \ldots) \triangleright m_n. \tag{3}$$

Therefore, multidimensional model (3) is specified by an ordered sequence of low-dimensional basic assignments — a *generating sequence* $m_1, m_2, \ldots, m_n$.

From the point of view of artificial intelligence models used to represent knowledge in a specific area of interest, a special role is played by the so-called *perfect sequences*, i.e., generating sequences $m_1, m_2, \ldots, m_n$, for which

$$m_1 \triangleright m_2 \triangleright \ldots \triangleright m_k = m_k \triangleright (m_1 \triangleright \ldots \triangleright m_{k-1})$$

for any $k = 2, \ldots, n$.

The property explaining why we call these sequences "perfect" is expressed by the following assertion proven in [6].

Theorem 4 *A generating sequence $m_1, m_2, \ldots, m_n$ is perfect if and only if all $m_1, m_2, \ldots, m_n$ are marginal assignments of the multidimensional assignment $m_1 \triangleright m_2 \triangleright \ldots \triangleright m_n$:*

$$(m_1 \triangleright m_2 \triangleright \ldots \triangleright m_n)^{\downarrow K_j} = m_j,$$

for all $j = 1, \ldots, n$.

3.3 Reconstruction of a belief network

Having a perfect sequence $m_1, m_2, \ldots, m_n$ (m_ℓ being the basic assignment of X_{K_ℓ}), we first order (in an arbitrary way) all the variables for which at least one of the basic assignments m_ℓ is defined. Finally we have

$$\{X_1, X_2, X_3, \ldots, X_k\} = \{X_i\}_{i \in K_1 \cup \ldots \cup K_n}.$$

Then we get a graph of the constructed belief network in the following way:

1. the nodes are all the variables $X_1, X_2, X_3, \ldots, X_k$;
2. there is an edge $(X_i \to X_j)$ if there exists a basic assignment m_ℓ such that both $i, j \in K_\ell$, $j \notin K_1 \cup \ldots \cup K_{\ell-1}$ and either $i \in K_1 \cup \ldots \cup K_{\ell-1}$ or $i < j$.

Evidently, for each j the requirement $j \in K_\ell$, $j \notin K_1 \cup \ldots \cup K_{\ell-1}$ is met exactly for one $\ell \in \{1, \ldots, n\}$. It means that all the parents of node X_j must be from the respective set $\{X_i\}_{i \in K_\ell}$ and therefore the necessary conditional belief function $Bel(X_j | X_{pa(j)})$ can easily be computed from basic assignment m_ℓ via (1) and some (not yet specified) conditioning rule. As far as we know, the use of a conditioning rule is still not fixed in evidence theory, and therefore we leave this question open for the present.

It is also evident, that if both i and j are in the same basic assignment and not in previous ones, then the direction of the arc depends only on the ordering of the variables. This might lead to different independences, nevertheless, the following theorem proven in [9] sets forth that any of them is induced by the perfect sequence.

Table 1: Basic assignments m_i and conditional basic assignments $m_{.|i}$.

$A \subseteq \mathbf{C}_i$	$m_i(A)$
$\{h_i\}$	0.49
$\{t_1\}$	0.49
$\{h_1, t_1\}$	0.02

$D \subseteq \mathbf{B}$	$m_{.\|i}(D)$
$\{b\}$	0.49
$\{\bar{b}\}$	0.51
$\{b, \bar{b}\}$	0.02

Table 2: Joint basic assignment m of variables C_1, C_2 and B.

m	$\{b\}$			$\{\bar{b}\}$			$\{b, \bar{b}\}$		
	$\{h_2\}$	$\{t_2\}$	$\{h_2, t_2\}$	$\{h_2\}$	$\{t_2\}$	$\{h_2, t_2\}$	$\{h_2\}$	$\{t_2\}$	$\{h_2, t_2\}$
$\{h_1\}$	0.2401	0	0	0	0.2401	0	0	0	0.0098
$\{t_1\}$	0	0.2401	0	0.2401	0	0	0	0	0.0098
$\{h_1, t_1\}$	0	0	0	0	0	0	0.0098	0.0098	0.0004

Theorem 5 *For a belief network defined by the above procedure the following independence statements are satisfied for any* $j = 2, \ldots k$*:*

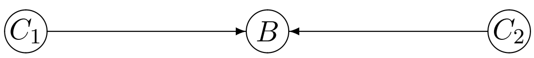
$$\{j\} \perp\!\!\!\perp (\{i < j\} \setminus pa(j)) \mid pa(j). \tag{4}$$

Let us note that it is different than in the case of directed evidential networks with conditional belief functions introduced in [3], where is no distinction between conditionally and unconditionally independent variables, as we shall see in the next section.

4 Example: two coins toss

Let us consider two fair coins toss expressed by variables C_1 and C_2 with values in $\mathbf{C}_1$ and $\mathbf{C}_2$, respectively ($\mathbf{C}_i = \{h_i, t_i\}$), and the basic assignments m_1 and m_2 (contained in the left part of Table 1) expressing the fact that the result of any of the coins may from time to time be unknown. The results of tossing two coins are usually considered to be independent, therefore the joint basic assignment m_{12} is just a product of these m_1 and m_2 (cf. definition of random set independence at the beginning of Section 2.3).

Now, let us consider one more variable B expressing the fact the bell is ringing, i.e $\mathbf{B} = \{b, \bar{b}\}$. It happens only if the result on both coins is the same (two heads or two tails). It is evident, that B depends on both C_1 and C_2, which corresponds to the following graph



and (due to deterministic dependence of the values of B on the values of C_1 and C_2) the joint basic assignment of the three variables is in Table 2. Above-presented graph can easily be obtained from perfect sequence of basic assignments m_1, m_2 and $m_3 \equiv m$ (contained in Tables 1 and 2) via the algorithm presented in the preceding section.

The approach suggested by Ben Yaghlane et al. [3] is completely different. The authors start from belief functions of C_1 and C_2 and conditional belief functions of B given C_1 and C_2, respectively. To make the difference between these two approaches more apparent we will use

Table 3: Joint belief function Bel of variables C_1, C_2 and B.

m	$\{b\}$			$\{\bar{b}\}$			$\{b, \bar{b}\}$		
	$\{h_2\}$	$\{t_2\}$	$\{h_2, t_2\}$	$\{h_2\}$	$\{t_2\}$	$\{h_2, t_2\}$	$\{h_2\}$	$\{t_2\}$	$\{h_2, t_2\}$
$\{h_1\}$	0.0624	0.0624	0.0025	0.0624	0.0624	0.0025	0.0001	0.0001	~ 0
$\{t_1\}$	0.0624	0.0624	0.0025	0.0624	0.0624	0.0025	0.0001	0.0001	~ 0
$\{h_1, t_1\}$	0.0025	0.0025	0.0001	0.0025	0.0025	0.0001	~ 0	~ 0	~ 0

basic assignments instead of belief functions (belief functions, nevertheless, can be easily obtained from them by (1)). The conditional basic assignments of B given C_1 and C_2, respectively, can be found in the right part of Table 1. Let us note that these conditional basic assignments do not depend on the condition, as the results of tossing two coins are independent and therefore also the event that the bell rings does not depend on the result at one coin.

The values of joint belief is computed from Tables 1 using (non-normalized) conjunctive combination rule. Results of these computations can be found in Table 3.

It is evident that the independence (non-interactivity) between coins C_1 and C_2 is not valid any more — it has been substituted by conditional non-interactivity, which does not make a sense, as C_1 is strongly dependent on C_2 whenever B is known.

References

[1] C. Beeri, R. Fagin, D. Maier, M. Yannakakis, On the desirability of acyclic database schemes, *J. of the Association for Computing Machinery,* **30** (1983), 479–513.

[2] B. Ben Yaghlane, Ph. Smets, K. Mellouli, Belief functions independence: II. the conditional case, *Int. J. Approx. Reasoning,* **31** (2002), 31–75.

[3] B. Ben Yaghlane, Ph. Smets, K. Mellouli, Directed evidential networks with conditional belief functions. In: Nielsen TD, Zhang NL (eds.) *Proceedings of ECSQARU 2003,* 291–305.

[4] S. Benferhat, D. Dubois, L. Gracia, H. Prade, Directed possibilistic graphs and possibilistic logic. In: Bouchon-Meunier B, Yager RR (eds.) *Proc. of IPMU'98,* 1470–1477.

[5] I. Couso, S. Moral, P. Walley, Examples of independence for imprecise probabilities. In: De Cooman G, Cozman FG, Moral S, Walley P (eds) *Proceedings ISIPTA'99,* 121–130.

[6] R. Jiroušek, J. Vejnarová, Compositional models and conditional independence in Evidence Theory, accepted to *Int. J. Approx. Reasoning.*

[7] P. P. Shenoy, Conditional independence in valuation-based systems. *Int. J. Approx. Reasoning,* **10** (1994) 203–234.

[8] J. Vejnarová, On conditional independence in evidence theory. In: *Proceedings of ISIPTA'09.* Durham, UK, 431–440.

[9] J. Vejnarová, An Alternative Approach to Evidential Network Construction. To appear in: *Proceedings of SMPS'10.* Oviedo, Spain.