An Alternative Approach to Evidential Network Construction

Jiřina Vejnarová

Abstract We present an alternative approach to belief network construction based on operator of composition of basic assignments. We show that belief networks constructed in this way have similar structural properties to Bayesian networks in contrary to previously proposed directed evidential networks by Ben Yaghlane at al.

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].

In this paper we bring an alternative to [3], which does not seem to us to be satisfactory, as graphical tools well-known from Bayesian networks are used in different sense. Our approach is based on previously introduced operator of composition for basic assignments [7, 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 is not solved here.

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 (and its basic properties) and in Section 4 after recalling perfect sequences of basic assignments we present an algorithm for transformation of a perfect sequence into an evidential network. We also demonstrate, through a simple example, in which sense our approach is superior to the previous one [3].

Institute of Information Theory and Automation of the ASCR, 182 08 Prague, Czech Republic, e-mail: vejnar@utia.cas.cz

2 Basic Notions

In this section we will briefly recall basic concepts from evidence theory [9] concerning sets, set functions and (conditional) independence.

2.1 Set Projections and Joins

For an index set $N = \{1,2,\dots,n\}$ 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* of $x = (x_1,x_2,\dots,x_n) \in \mathbf{X}_N$ into $\mathbf{X}_K$ will be denoted $x^{\downarrow K}$, i.e., for $K = \{i_1,i_2,\dots,i_k\}$

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

Analogously, for $M \subset K \subseteq N$ and $A \subset \mathbf{X}_K$, $A^{\downarrow M}$ will denote a *projection* of A into $\mathbf{X}_M$:[1]

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

In addition to the projection, in this text we will also need an opposite operation, which will be called a join. By a *join*[2] 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 [9] (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.,

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

where $\mathscr{P}(\mathbf{X}_N)$ is power set of $\mathbf{X}_N$ and $\sum_{A\subseteq \mathbf{X}_N} m(A) = 1$. Furthermore, we assume that $m(\emptyset) = 0$. A set $A \in \mathscr{P}(\mathbf{X}_N)$ is a *focal element* if $m(A) > 0$.

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

---

1 Let us remark that we do not exclude situations when $M = \emptyset$. In this case $A^{\downarrow\emptyset} = \emptyset$.

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

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

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 $K \cap L = \emptyset$.

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 [10, 2, 11]; 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 [11, 6], as we consider it more suitable (the arguments can be found in [11]).

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

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

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 [11] that this conditional independence concept satisfies so-called semi-graphoid properties taken as reasonable to be valid for any conditional independence concept (see e.g. [8]).

3 Operator of Composition and Its Basic Properties

Operator of composition of basic assignments was introduced in [7] 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

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

Its basic properties are contained in the following lemma proven in [7].

Lemma 1. *For arbitrary two basic assignments m_1 on $\mathbf{X}_K$ and m_2 on $\mathbf{X}_L$ the following properties hold true:*

(i) $m_1 \triangleright m_2$ *is a basic assignment on* $\mathbf{X}_{K\cup L}$*,*
(ii) $(m_1 \triangleright m_2)^{\downarrow K} = m_1$*,*
(iii) $m_1 \triangleright m_2 = m_2 \triangleright m_1 \iff m_1^{\downarrow K\cap L} = m_2^{\downarrow K\cap L}$.

From these basic properties one can see that operator of composition is not commutative in general, but it preserves first marginal (in case of projective basic assignments both of them). In both these aspects it differs from conjunctive combination rule. Furthermore, operator of composition is not associative and therefore its iterative applications must be made carefully, as we will see in the next section.

A lot of other properties possessed by the operator of composition can be found in [6, 7], nevertheless here we will confine ourselves to the following theorem (proven in [6]) expressing the relationship between conditional independence and operator of composition.

Theorem 1. *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 $A \subseteq \mathbf{X}_{K\cup L}$.

4 Belief Network Generated by a Perfect Sequence

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 $\mathbf{X}_{K_1 \cup K_2 \cup \ldots \cup K_n}$. However, since we know that the operator of composition is neither commutative nor associative, we have to properly specify what "composing them together" means.

To avoid using too many parentheses let us make the following convention. Whenever we put down the expression $m_1 \triangleright m_2 \triangleright \ldots \triangleright m_n$ we will understand that the operator of composition is performed successively from left to right:[3]

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

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

4.1 Perfect Sequences

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

$$\begin{aligned} m_1 \triangleright m_2 &= m_2 \triangleright m_1, \\ m_1 \triangleright m_2 \triangleright m_3 &= m_3 \triangleright (m_1 \triangleright m_2), \\ &\vdots \\ m_1 \triangleright m_2 \triangleright \ldots \triangleright m_n &= m_n \triangleright (m_1 \triangleright \ldots \triangleright m_{n-1}). \end{aligned}$$

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

Theorem 2. *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] Naturally, if we want to change the ordering in which the operators are to be performed we will do so using parentheses.

4.2 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 all the variables for which at least one of the basic assignments m_ℓ is defined in such a way that first we order (in an arbitrary way) variables for which m_1 is defined, then variables from m_2 which are not contained in m_1, etc.[4] 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 $(X_i \rightarrow X_j)$ 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 sets forth that any of them is induced by the perfect sequence.

Theorem 3. *For a belief network defined by the above procedure the following independence statements are satisfied for any $j = 2, \ldots k$:*

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

Proof. Let $j \in K_\ell$, $j \notin K_1 \cup \ldots \cup K_{\ell-1}$. Due to the fact that

$$m_1 \triangleright m_2 \triangleright \ldots \triangleright m_{\ell-1} \triangleright m_\ell = (\cdots(m_1 \triangleright m_2) \triangleright \cdots \triangleright m_{\ell-1}) \triangleright m_\ell$$

and Theorem 1 we have that

$$K_\ell \setminus (K_1 \cup \ldots \cup K_{\ell-1}) \perp\!\!\!\perp (K_1 \cup \ldots \cup K_{\ell-1}) \setminus K_\ell \,|\, K_\ell \cap (K_1 \cup \ldots \cup K_{\ell-1}). \tag{5}$$

It is evident that $(K_1 \cup \ldots \cup K_{\ell-1}) \setminus K_\ell = \{i < j\} \setminus pa(j)$, let us denote it by L. Now, there are two possibilities: either $K_\ell \cap (K_1 \cup \ldots \cup K_{\ell-1}) = pa(j)$ (if j does not have any parents appearing first in K_ℓ) or $K_\ell \cap (K_1 \cup \ldots \cup K_{\ell-1}) \subsetneq pa(j)$ (otherwise).

---

[4] Let us note that variables $X_1, X_2, \ldots, X_k$ may be ordered arbitrarily, nevertheless, for the above ordering proof of Theorem 3 is simpler than in the general case.

In the first case either $K_\ell \setminus (K_1 \cup \ldots \cup K_{\ell-1}) = \{j\}$ and we immediately obtain (4), or $K_\ell \setminus (K_1 \cup \ldots \cup K_{\ell-1}) \supsetneq \{j\}$ and (4) follows from (5) due to $K \cup M \perp\!\!\!\perp L|I\ [m] \Rightarrow K \perp\!\!\!\perp L|I\ [m]$ (following for any mutually disjoint sets I, K, L, M from semi-graphoid properties), where $K = \{j\}, M = K_\ell \setminus (K_1 \cup \ldots \cup K_{\ell-1}) \setminus \{j\}$ and $I = K_\ell \cap (K_1 \cup \ldots \cup K_{\ell-1}) = pa(j)$.

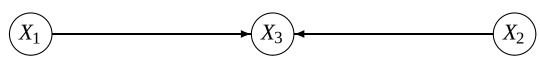
In the latter case, we start by application of the implication $K \cup M \perp\!\!\!\perp L|I\ [m] \Rightarrow K \perp\!\!\!\perp L|M \cup I\ [m]$, whose validity for any mutually disjoint sets I, K, L, M follows again from semi-graphoid properties, to $K = K_\ell \setminus (K_1 \cup \ldots \cup K_{\ell-1}) \setminus \{j\} \setminus pa(j)$, $M = K_\ell \setminus (K_1 \cup \ldots \cup K_{\ell-1}) \cap pa(j)$ and $I = K_\ell \cap (K_1 \cup \ldots \cup K_{\ell-1})$. As $M \cup I = \{i < j\} \setminus pa(j)$ we can then proceed analogous to previous paragraph to obtain (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 the following simple example suggests.

Example 1. Let us consider a sequence of basic assignments m_1, m_2 and m_3, defined on $\mathbf{X}_1, \mathbf{X}_2$ and $\mathbf{X}_1 \times \mathbf{X}_2 \times \mathbf{X}_3$. This sequence need not be perfect, in general, but it is perfect iff

$$m_3^{\downarrow\{1,2\}}(x_1, x_2) = m_1(x_1) \cdot m_2(x_2).$$

This perfect sequence induces independence statements $1 \perp\!\!\!\perp 2$, but generally not $1 \perp\!\!\!\perp 2|3$. Using the above-presented algorithm, we can easily obtain the following graph expressing the same independence statements.



On the other hand, in [3] the same situation is described by $Bel(X_1)$, $Bel(X_2)$, $Bel(X_3|X_1)$ and $Bel(X_3|X_2)$ and the joint belief function is computed using conjunctive combination rule. Therefore, in the resulting model $X_1 \perp_m X_2|X_3$, which corresponds rather to so-called pseudobayesian networks than to Bayesian ones.

5 Conclusions

We introduced an alternative approach to evidential network construction to that presented in [3]. The evidential network is constructed from so-called perfect sequences of basic assignments through a simple transformation algorithm. We proved that the independence relations in the resulting models are analogous to those valid in Bayesian networks, while it does not hold for models introduced in [3]. Due to the limited extent of the paper we are not able to bring more detailed comparison of these two approaches, but we believe that Theorem 3 and Example 1 give the basic idea. Nevertheless, still one substantial problem should be solved — the choice of a proper conditioning rule compatible with (conditional) independence concept used in our models. It will be one of the main goals of our future research.

Acknowledgements 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.

References

1. Beeri, C., Fagin, R., Maier, D., Yannakakis, M.: On the desirability of acyclic database schemes. J. Assoc. Comput. Mach. 30, 479–513 (1983)
2. Ben Yaghlane, B., Smets, Ph., Mellouli, K.: Belief functions independence: II. the conditional case. Internat. J. Approx. Reason. 31, 31–75 (2002)
3. Ben Yaghlane, B., Smets, Ph., Mellouli, K.: Directed evidential networks with conditional belief functions. In: Nielsen, T.D., Zhang, N.L. (eds.) Proceedings of the 7th European Conference Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU 2003, Aalborg, Denmark). Lecture Notes in Computer Science, vol. 2711, pp. 291–305 (2003)
4. Benferhat, S., Dubois, D., Gracia, L., Prade, H.: Directed possibilistic graphs and possibilistic logic. In: Bouchon-Meunier, B., Yager, R.R. (eds.) Proceedings of the 7th International Conference on Information Processing and Management of Uncertainty in Knowledge-based Systems (IPMU'98, Paris, France) pp. 1470–1477 (1998)
5. Couso, I., Moral, S., Walley, P.: Examples of independence for imprecise probabilities. In: de Cooman, G., Cozman, F.G., Moral, S., Walley, P. (eds.) Proceedings of the First International Symposium on Imprecise Probabilities and Their Applications (ISIPTA'99, Ghent, Belgium) pp. 121–130 (1999)
6. Jiroušek, R., Vejnarová, J.: Compositional models and conditional independence in evidence theory. Int. J. Approx. Reasoning (2010) doi: 10.1016/j.ijar.2010.02.005
7. Jiroušek, R., Vejnarová, J., Daniel, M.: Compositional models for belief functions. In: de Cooman, G., Vejnarová, J., Zaffalon, M. (eds) Proceedings of the Fifth International Symposium on Imprecise Probability: Theories and Applications (ISIPTA'07, Prague, Czech Republic), pp. 243–252 (2007)
8. Lauritzen, S.L.: Graphical models. Oxford University Press, Oxford (1996)
9. Shafer, G.: A Mathematical Theory of Evidence. Princeton University Press, Princeton, New Jersey (1976)
10. Shenoy, P.P.: Conditional independence in valuation-based systems. Int. J. Approx. Reasoning, 10, 203–234 (1994)
11. Vejnarová, J.: On conditional independence in evidence theory. In: Augustin, T., Coolen, F.P.A., Moral, S., Troffaes, M.C.M. (eds.) Proceedings of the Sixth International Symposium on Imprecise Probability: Theories and Applications (ISIPTA'09, Durham, UK) pp. 431–440 (2009)