ADVANCES IN INTELLIGENT AND *SOFT COMPUTING* 77

Christian Borgelt
Gil González-Rodríguez
Wolfgang Trutschnig
María Asunción Lubiano
María Ángeles Gil
Przemysław Grzegorzewski
Olgierd Hryniewicz (Eds.)

Combining Soft Computing and Statistical Methods in Data Analysis

Springer

cajAstur

Editors

Christian Borgelt
Research Unit on Intelligent
Data Analysis and Graphical Models
European Centre for Soft Computing
Edificio Científico-Tecnológico. 3ª Planta
C/ Gonzalo Gutiérrez Quirós s/n
33600 Mieres, Spain
E-mail: christian.borgelt@softcomputing.es

Gil González-Rodríguez
Research Unit on Intelligent
Data Analysis and Graphical Models
European Centre for Soft Computing
Edificio Científico-Tecnológico. 3ª Planta
C/ Gonzalo Gutiérrez Quirós s/n
33600 Mieres, Spain
E-mail: gil.gonzalez@softcomputing.es

Wolfgang Trutschnig
Research Unit on Intelligent
Data Analysis and Graphical Models
European Centre for Soft Computing
Edificio Científico-Tecnológico. 3ª Planta
C/ Gonzalo Gutiérrez Quirós s/n
33600 Mieres, Spain
E-mail: wolfgang.trutschnig@softcomputing.es

María Asunción Lubiano
Departamento de Estadística e I.O. y D.M.
Universidad de Oviedo, Facultad de Ciencias
C/ Calvo Sotelo s/n, 33007 Oviedo, Spain
E-mail: lubiano@uniovi.es

María Ángeles Gil
Departamento de Estadística
e I.O. y D.M.
Universidad de Oviedo
Facultad de Ciencias
C/ Calvo Sotelo s/n
33007 Oviedo, Spain
E-mail: magil@uniovi.es

Dr. Przemysław Grzegorzewski
Systems Research Institute
Polish Academy of Sciences
Newelska 6, 01-447 Warsaw,
Poland

and

Faculty of Mathematics and
Information Science
Warsaw University of Technology
Plac Politechniki 1, 00-661 Warsaw
Poland
E-mail: pgrzeg@ibspan.waw.pl

Prof. Dr. Olgierd Hryniewicz
Systems Research Institute
Polish Academy of Science
Newelska 6, 01-447 Warsaw
Poland
E-mail: hryniewi@ibspan.waw.pl

ISBN 978-3-642-14745-6 e-ISBN 978-3-642-14746-3

DOI 10.1007/978-3-642-14746-3

Advances in Intelligent and Soft Computing ISSN 1867-5662

Library of Congress Control Number: 2010931605

Typeset & Cover Design: Scientific Publishing Services Pvt. Ltd., Chennai, India.

Printed on acid-free paper

5 4 3 2 1 0

springer.com

An Attempt to Define Graphical Models in Dempster-Shafer Theory of Evidence

Radim Jiroušek

Abstract. The goal of this paper is to introduce graphical models in Dempster-Shafer theory of evidence. The way the models are defined is a natural and straightforward generalization of the approach from probability theory. The models possess the same "Global Markov Properties", which holds for probabilistic graphical models. Nevertheless, the last statement is true only under the assumption that one accepts a new definition of conditional independence in Dempster-Shafer theory, which was introduced in Jiroušek and Vejnarová (2010). Therefore, one can consider this paper as an additional reason supporting this new type of definition.

Keywords: Graphical Markov models, Conditional independence, Factorization, Multidimensional basic assignment.

1 Introduction

Graphical Markov models [8] developed to their variety and proficiency in the last two decades of the 20th century, have become a benchmark with which models from other theories of uncertainty are often compared. Here we have in mind Bayesian networks (perhaps the most popular member of graphical Markov models), decomposable models (indisputably the most efficient from the computational point of view) and also "classical" graphical models. The last models were originally studied within the class of log-linear models as distributions whose interactions can be described with the help of simple graphs.

Radim Jiroušek
Faculty of Management, University of Economics, Jindřichuv Hradec, and Institute of Information Theory and Automation, Academy of Sciences,
Prage, Czech Republic
e-mail: `radim@utia.cas.cz`

C. Borgelt et al. (Eds.): Combining Soft Computing & Stats. Methods, AISC 77, pp. 361–368.
springerlink.com

In this paper we want to show that the idea upon which graphical models were founded can be (almost straightforwardly) exploit also within Dempster-Shafer Theory of evidence. In fact, the only new idea of the approach is that not all subsets of the considered space of discernment may be focal elements.

2 Basic Concepts and Notation

In the following text we will need just basic concepts od Dempster-Shafer theory of evidence. However, to make the explanation more lucid we will explain our motivation originated in probability theory. Naturally, when speaking about graphical models we cannot avoid a couple of notions from graph theory. All these concepts will be briefly introduced in this section.

All our considerations will concern finite multidimensional space

$$\mathbf{X}_N = \mathbf{X}_1 \times \mathbf{X}_2 \times \ldots \times \mathbf{X}_n. \quad (1)$$

The reader can interpret it either as a space of possible combinations of values of n (random) variables, or as an n-dimensional space on which the respective measures will be defined. Subsets of $N = \{1, 2, \ldots, n\}$ will be denoted by K, L, M with possible indices. So, $\mathbf{X}_K$ will denote a Cartesian product of those $\mathbf{X}_i$, for which $i \in K$:

$$\mathbf{X}_K = \times_{i \in K} \mathbf{X}_i.$$

A *projection* of $x = (x_1, x_2, \ldots, x_n) \in \mathbf{X}_N$ into $\mathbf{X}_K$ will be denoted $x^{\downarrow K}$, i.e. for $K = \{i_1, i_2, \ldots, i_\ell\}$

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

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

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

Let us remark that we do not exclude situations when $K = \emptyset$: $A^{\downarrow \emptyset} = \emptyset$.

One of the most important notions of this text will be a *join* of two subsets $A \subseteq \mathbf{X}_K$ and $B \subseteq \mathbf{X}_L$, which is defined

$$A \otimes B = \{x \in \mathbf{X}_{K \cup L} : x^{\downarrow K} \in A \;\; \& \;\; x^{\downarrow L} \in B\}. \quad (2)$$

Notice that if K and L are disjoint then the join of the corresponding sets is just their Cartesian product $A \otimes B = A \times B$. For $K = L$, $A \otimes B = A \cap B$. If $K \cap L \neq \emptyset$ and $A^{\downarrow K \cap L} \cap B^{\downarrow K \cap L} = \emptyset$ then also $A \otimes B = \emptyset$.

In view of this paper it is important to realize that if $x \in C \subseteq \mathbf{X}_{K \cup L}$, then $x^{\downarrow K} \in C^{\downarrow K}$ and $x^{\downarrow L} \in C^{\downarrow L}$, which means that always $C \subseteq C^{\downarrow K} \otimes C^{\downarrow L}$. However, it does not mean that $C = C^{\downarrow K} \otimes C^{\downarrow L}$. For example, considering only a 2-dimensional frame of discernment $\mathbf{X}_{\{1,2\}}$ with $\mathbf{X}_i = \{a_i, \bar{a}_i\}$ for both $i = 1, 2$, and $C = \{a_1 a_2, \bar{a}_1 a_2, a_1 \bar{a}_2\}$ one gets

$$C^{\downarrow \{1\}} \otimes C^{\downarrow \{2\}} = \{a_1, \bar{a}_1\} \otimes \{a_2, \bar{a}_2\} = \{a_1 a_2, \bar{a}_1 a_2, a_1 \bar{a}_2, \bar{a}_1 \bar{a}_2\} \neq C.$$

2.1 Graph Notions

In the paper we will exclusively consider simple graphs $G = (N,E)$ with a set of nodes N corresponding to the previously introduced index set. It means that the considered graphs contain neither oriented nor multiple edges and also no loops.

An important notion is that of a *clique*, which denotes a maximal subset of N inducing a complete subgraph (i.e. all pairs of nodes of a clique are connected by an edge and adding an additional node to the clique violates this property). The graph in Figure 1(a) has three cliques: $\{1,2,3,4\},\{3,4,5\},\{6\}$, the graph in Figure 1(b) has five cliques: $\{1,2,3\},\{1,4\},\{3,6\},\{4,5\},\{5,6\}$.

A graph is *decomposable* if its cliques $K_1,K_2,\ldots,K_r$ can be ordered in the way that the sequence meets the so called *running intersection property* (RIP):

$$\forall i=2,\ldots,r \quad \exists j(1 \leq j < i): \quad K_i \cap (K_1 \cup \ldots \cup K_{i-1}) \subseteq K_j. \tag{3}$$

Notice that this property is met by any ordering of the cliques of the graph in Figure 1(a), and that the cliques of the graph in Figure 1(b) cannot be ordered to meet this property. It means that from the mentioned two graphs only the former is decomposable. The graph in Figure 1(c) is also decomposable, because the ordering of its cliques $\{1,2,4\},\{2,3,4\},\{4,6\},\{3,4,5\}$ meets RIP (in spite of the fact that, for example, $\{3,4,5\},\{1,2,4\},\{2,3,4\},\{4,6\}$ does not meet this property).

The last notions we will need are notions of *separation* and a *separating set.* We say that two different nodes $i,j \in N$ are *separated by a set* $K \subseteq N \setminus \{i,j\}$ if we cannot go along the graph edges from i to j without going through a node from K. So, if there is no path from i to j (as, for example there is no path from 1 to 6 in the graph in Figure 1(a)) then even the empty set may be a separating set. A set K is a *minimal separating set* if there exists a pair of nodes i and j, which is separated by K but no proper subset of K separates i and j. Notice that in the graph in Figure 1(c) both $\{2,4\}$ and $\{4\}$ are minimal separating sets; the former is a minimal separating set for 1 and 3, whereas the latter is a minimal separating set for 1 and 6.

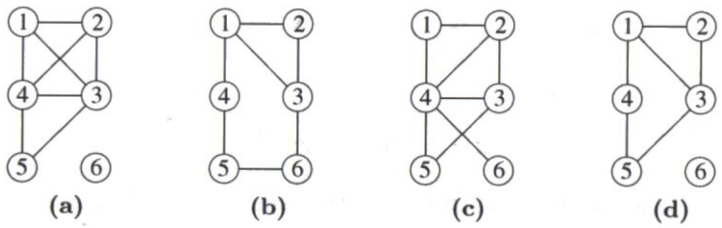

Fig. 1 Graphs with 6 nodes

If graph $G=(N,E)$ is not complete then it is always possible to find a couple of subsets $L,M \subset N$ (usually there are lot of such couples; the exception is a graph consisting of only two cliques, for which this couple is unique) such that

- $L \cup M = N$;
- $L \cap M$ is a minimal separating set;
- each pair of nodes $i \in L \setminus M$, $j \in M \setminus L$ is separated by $L \cap M$.

The set of all these couples will be denoted by symbol $\mathscr{S}(G)$ - for examples concerning all the graphs in Figure 1 see Table 1. Now, we are ready to introduce a class of subsets of $\mathbf{X}_N$ whose structures *comply with graph* G (these sets will be used in the definition of graphical models in Section 4):

$$\mathscr{R}(G) = \{A \subseteq \mathbf{X}_N : \forall (L,M) \in \mathscr{S}(G) \ \ (A = A^{\downarrow L} \otimes A^{\downarrow M})\}. \tag{4}$$

Table 1 $\mathscr{S}(G)$ for graphs in Figure 1

Graph G	Couples (L,M) from $\mathscr{S}(G)$
(a)	$(\{1,2,3,4\},\{3,4,5,6\})$
	$(\{1,2,3,4,6\},\{3,4,5\})$
	$(\{1,2,3,4,5\},\{6\})$
(d)	$(\{1,2,3,4,5\},\{6\})$
	$(\{1,2,3\},\{1,3,4,5,6\})$
	$(\{1,2,3,6\},\{1,3,4,5\})$
	$(\{1,2,3,5,6\},\{1,4,5\})$
	$(\{1,2,3,5\},\{1,4,5,6\})$
	$(\{1,2,3,4\},\{3,4,5,6\})$
	$(\{1,2,3,4,6\},\{3,4,5\})$
(b)	$(\{1,2,4\},\{2,3,4,5,6\})$
	$(\{1,2,3,4\},\{3,4,5,6\})$
	$(\{1,2,3,4,5\},\{3,5,6\})$
	$(\{1,2,3,4,6\},\{4,5,6\})$
(c)	$(\{1,2,4\},\{2,3,4,5,6\})$
	$(\{1,2,4,6\},\{2,3,4,5\})$
	$(\{1,2,3,4\},\{3,4,5,6\})$
	$(\{1,2,3,4,6\},\{3,4,5\})$
	$(\{1,2,3,4,5\},\{4,6\})$

2.2 Probabilistic Factorization

Consider a probability measure π on $\mathbf{X}_N$ and $L,M \subseteq N$ such that $L \cup M = N$. We say that π factorizes with respect to a couple (L,M) if the exist functions

$$\phi : \mathbf{X}_L \longrightarrow [0,+\infty), \qquad \psi : \mathbf{X}_M \longrightarrow [0,+\infty),$$

such that for all $x \in \mathbf{X}_N$

$$\pi(x) = \phi(x^{\downarrow L}) \cdot \psi(x^{\downarrow M}).$$

It is well known that π factorizes with respect to (L,M) if and only if for all $x \in \mathbf{X}_N$

$$\pi(x) \cdot \pi^{\downarrow L \cap M}(x^{\downarrow L \cap M}) = \pi^{\downarrow L}(x^{\downarrow L}) \cdot \pi^{\downarrow M}(x^{\downarrow M}),$$

which corresponds to the conditional independence $L \setminus M \perp\!\!\!\perp M \setminus L | L \cap M \, [\pi]$.

This notion forms a basis for a more general notion of a *graphical model*, which is a probability distribution factorizing with respect to a graph $G = (N,E)$ [8].

Consider a graph $G = (N,E)$ with r cliques $K_1, K_2, \ldots, K_r$. We say that a probability distribution π *factorizes with respect to graph* G if there exist r functions $\phi_1, \phi_2, \ldots, \phi_r$,

$$\phi_i : \mathbf{X}_{K_i} \longrightarrow [0, +\infty),$$

such that for all $x \in \mathbf{X}_N$

$$\pi(x) = \prod_{i=1}^{r} \phi_i(x^{\downarrow K_i}).$$

What is the advantage of graphical models? Naturally, first of all we can represent such a distribution with the help of (in the binary case) $\prod_{i=1}^{r} 2^{|K_i|}$ parameters (factors), which is usually much less than 2^n, the number of probabilities necessary to define a general n-dimensional distribution. Moreover, graphical models have their "semantics" expressible with the help of their conditional independence structure: If distribution π factorizes with respect to $G = (N,E)$ and $K \subset N$ separates in G nodes $i, j \in N$, then $i \perp\!\!\!\perp j \,|\, K \,[\pi]$.

2.3 Basic Assignment Notation

The role of a probability distribution from a probability theory is in Dempster-Shafer theory played by any of the set functions: belief function, plausibility function, commonality function or basic (*probability or belief*) assignment [4, 9]. In this text we will exclusively use normalized basic assignments for the purpose. Such a *basic assignment* m on $\mathbf{X}_K$ $(K \subseteq N)$ is a function

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

for which $m(\emptyset) = 0$, and $\sum_{A \subseteq \mathbf{X}_K} m(A) = 1$. All the sets A for which $m(A)$ is positive are called *focal elements* of m.

Having a basic assignment m on $\mathbf{X}_K$ we will consider its *marginal assignment* on $\mathbf{X}_L$ (for $L \subseteq K$), which is defined (for each $\emptyset \neq B \subseteq \mathbf{X}_L$):

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

3 Factorization and Independence

Unconditional (*marginal*) *independence* has been introduced in Dempster-Shafer theory in several equivalent ways; mostly as an application of *conjunctive combination rule* (non-normalized *Dempster's rule* of combination) [1, 3, 7, 10], or with the help of *commonality functions* [12, 11]. Here, we will use another (and as it was showed in [6] still equivalent) definition.

Let $K, L \subset N$ be disjoint. For a basic assignment m the independence $K \perp\!\!\!\perp L[m]$ holds if for all $A \subseteq \mathbf{X}_{K \cup L}$

$$m^{\downarrow K \cup L} = \begin{cases} m^{\downarrow K} \cdot m^{\downarrow L} & \text{if } A = A^{\downarrow K} \otimes A^{\downarrow L} \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

Having a probability measure π defined on a 2-dimensional space $\mathbf{X}_1 \times \mathbf{X}_2$ and factorizing with respect to $(\{1\},\{2\})$ we know that there exist functions ϕ and ψ such that for each $x \in \mathbf{X}_1 \times \mathbf{X}_2$

$$\pi(x) = \phi(x^{\downarrow \{1\}}) \cdot \psi(x^{\downarrow \{2\}}). \tag{6}$$

It means that $|\mathbf{X}_1| \cdot |\mathbf{X}_2|$ probabilities of measure π is defined with the help of $|\mathbf{X}_1|$ and $|\mathbf{X}_2|$ values of the factor functions ϕ and ψ. This fits the product rule expressed by formula (6).

Is it possible to transfer this simple idea directly into Dempster-Shafer theory? Basic assignment m on $\mathbf{X}_1 \times \mathbf{X}_2$ is defined with the help of $2^{|\mathbf{X}_1| \cdot |\mathbf{X}_2|}$ values, whereas factor functions

$$\mu : \mathscr{P}(\mathbf{X}_1) \longrightarrow [0, +\infty), \qquad \nu : \mathscr{P}(\mathbf{X}_2) \longrightarrow [0, +\infty),$$

are defined with the help of $2^{|\mathbf{X}_1|}$ and $2^{|\mathbf{X}_2|}$ values, respectively. Thus using an analogy to a product rule we can get only $2^{|\mathbf{X}_1| + |\mathbf{X}_2|}$ different values. However noticing that factorization (in this simple 2-dimensional situation) should yield the independence $\{1\} \perp\!\!\!\perp \{2\}$, and looking at the definition formula (5), we see that we do not need to define values of m for all subset $A \subseteq \mathbf{X}_1 \times \mathbf{X}_2$, but only for those A for which $A = A^{\downarrow |1|} \otimes A^{\downarrow |2|}$.

Generalizing the above consideration to a more complex, overlapping factorization we proposed the following definition of factorization in [5].

Definition 1. ***Simple Factorization.*** Consider two nonempty sets $K \cup L = N$. We say that basic assignment m *factorizes with respect to* (K, L) if there exist two nonnegative set functions

$$\mu : \mathscr{P}(\mathbf{X}_K) \longrightarrow [0, +\infty), \quad \nu : \mathscr{P}(\mathbf{X}_L) \longrightarrow [0, +\infty),$$

such that for all $A \subseteq \mathbf{X}_{K \cup L}$

$$m(A) = \begin{cases} \phi(A^{\downarrow K}) \cdot \psi(A^{\downarrow L}) & \text{if } A = A^{\downarrow K} \otimes A^{\downarrow L} \\ 0 & \text{otherwise.} \end{cases}$$

It is almost obvious that for this notion the following simplified version of Factorization Lemma is valid [13].

Lemma 1. *Let $K, L \subseteq N$ be disjoint and nonempty, $K \cup L = N$. m factorizes with respect to (K, L) if and only if $K \setminus L \perp\!\!\!\perp L \setminus K | K \cap L$ $[m]$.*

4 Graphical Models

Definition 2. *Let $G = (N,E)$ be a graph with r cliques $K_1, K_2, \ldots, K_r$. We say that basic assignment m* factorizes with respect to graph G *if there exist r functions $\mu_1, \mu_2, \ldots, \mu_r$, $(\mu_i : \mathscr{P}(\mathbf{X}_{K_i}) \longrightarrow [0,+\infty))$, such that for all $A \subseteq \mathbf{X}_N$*

$$m(A) = \begin{cases} \prod_{i=1}^{r} \mu_i(A^{\downarrow K_i}), & \text{if } A \in \mathscr{R}(G), \\ 0 & \text{otherwise.} \end{cases}$$

Example 1. Consider a 6-dimensional basic assignment factorizing with respect to the graph in Figure 1(d). If all $\mathbf{X}_i$ are binary, then general basic assignment may have up to $2^{64} - 1$ focal elements. Nevertheless, since the considered graph consists of 5 cliques: $\{1,2,3\}$, $\{1,4\}$, $\{3,5\}$, $\{4,5\}$ and $\{6\}$, all the necessary factor functions are defined with by $2^8 + 3 \cdot 2^4 + 2^2 = 308$ numbers.

We believe that the above presented example sufficiently illustrates an efficiency with which graphical models can be represented in Dempster-Shafer theory. What remains to be showed that it possesses also the second advantageous property of probabilistic graphical models, i.e. that the dependence structure of the distribution is somehow encoded in the graph. We do not have enough space to formalize the property in a form of a theorem and to prove it but an analogy of the probabilistic statement presented at the end of Section 2.2 holds: If basic assignment m factorizes with respect to $G = (N,E)$ and $K \subset N$ separates nodes i, j in G, then $i \perp\!\!\!\perp j \,|\, K\,[m]$.

For this, however, we have to say what we understand by conditional independence in Dempster-Shafer theory. Namely, we cannot apply the definition used by most of the other authors (e.g. [2, 10, 12]) but the following definition introduced in [6].

Definition 3. ***Conditional Independence.*** Let $K, L, M \subset N$ be disjoint, K, L nonempty. We say that for a basic assignment m conditional independence $K \perp\!\!\!\perp L|M\ [m]$ holds if for any $A \subseteq \mathbf{X}_{K \cup L \cup M}$ such that $A = A^{\downarrow K \cup M} \otimes A^{\downarrow L \cup M}$ the equality

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

holds, and $m^{\downarrow K \cup L \cup M}(A) = 0$ for all the remaining $A \subseteq \mathbf{X}_{K \cup L \cup M}$, for which $A \neq A^{\downarrow K \cup M} \otimes A^{\downarrow L \cup M}$.

5 Conclusions

We have introduced graphical models in Dempster-Shafer theory as a simple and natural generalization of probabilistic graphical models. Analogously to probabilistic case, also for Dempster-Shafer graphical models one can show that they can be efficiently represented with a reasonable number of parameters and that some conditional independence relations can be read from the respective graphs. This holds, however, only when a new definition of conditional independence in Dempster-Shafer theory (see Definition 3) is accepted. Thus the paper brings an additional reason supporting this new definition. Recall that the new concept of conditional independence does not suffer from *inconsistency with marginalization* (for details and a Studený's example see [2]), for Bayesian basic assignments coincides with probabilistic conditional independence, and meets all the semigraphoid axioms.

Acknowledgements. This work was supported by National Science Foundation of the Czech Republic under the grants no. ICC/08/E010, and 201/09/1891, and by Ministry of Education under the grants 1M0572 and 2C06019.

References

1. Ben Yaghlane, B., Smets, P., Mellouli, K.: Belief Function Independence: I. The Marginal Case. Internat. J. Approx. Reason. 29, 47–70 (2002)
2. Ben Yaghlane, B., Smets, P., Mellouli, K.: Belief Function Independence: II. The Conditional Case. Internat. J. Approx. Reason. 31, 31–75 (2002)
3. Couso, I., Moral, S., Walley, P.: Examples of independence for imprecise probabilities. In: Proceedings of the First International Symposium on Imprecise Probabilities and Their Applications, ISIPTA 1999, Ghent, Belgium, pp. 121–130 (1999)
4. Dempster, A.: Upper and lower probabilities induced by a multi-valued mapping. Ann. Math. Statist. 38, 325–339 (1967)
5. Jiroušek, R.: Factorization and Decomposable Models in Dempster-Shafer Theory of Evidence. In: Workshop on the Theory of Belief Functions, Brest, France (2010)
6. Jiroušek, R., Vejnarová, J.: Compositional models and conditional independence in evidence theory. Internat. J. Approx. Reason (2010), doi:10.1016/j.ijar.2010.02.005
7. Klir, G.J.: Uncertainty and Information. In: Foundations of Generalized Information Theory. Wiley, Hoboken (2006)
8. Lauritzen, S.L.: Graphical models. Oxford University Press, Oxford (1996)
9. Shafer, G.: A Mathematical Theory of Evidence. Princeton University Press, New Jersey (1976)
10. Shenoy, P.P.: Conditional independence in valuation-based systems. Internat. J. Approx. Reason. 10, 203–234 (1994)
11. Studený, M.: Formal properties of conditional independence in different calculi of AI. In: Moral, S., Kruse, R., Clarke, E. (eds.) ECSQARU 1993. LNCS, vol. 747, pp. 341–351. Springer, Heidelberg (1993)
12. Studený, M.: On stochastic conditional independence: the problems of characterization and description. Ann. Math. Artif. Intell. 35, 323–341 (2002)
13. Vejnarová, J.: On conditional independence in evidence theory. In: Proceedings of the Sixth International Symposium on Imprecise Probability: Theories and Applications, ISIPTA 2009, Durham, UK, pp. 431–440 (2009)