Formal Properties of Conditional Independence in Different Calculi of AI[1]

Milan Studený

Institute of Information Theory and Automation
Academy of Sciences of Czech Republic
Pod vodárenskou věží 4, 18208 Prague 8, Czech Republic
e-mail: studeny@cspgas11.bitnet

Abstract. In this paper formal properties of CI in different frameworks are studied. The first part is devoted to the comparison of three different frameworks for study CI: probability theory, theory of relational databases and Spohn's theory of ordinal conditional functions. Although CI–models arising in these frameworks are very similar (they satisfy semigraphoid axioms) we give examples showing that their formal properties still differ (each other). On the other hand, we find that (within each of these frameworks) there exists no finite complete axiomatic characterization of CI–models by finding an infinite set of sound inference rules (the same in all three frameworks). In the second part further frameworks for CI are discussed: Dempster–Shafer theory, possibility theory and (general) Shenoy's theory of valuation–based systems.

1 Introduction

The concept of *conditional independence* (CI) seems to attract attention of researches in last two decades. The properties of CI were studied in several branches of AI, let us mention some of them :

- probabilistic reasoning
- theory of relational databases
- theory of ordinal conditional functions
- theory of belief functions
- possibility theory.

The reason is evident – the knowledge concerning independence (resp. dependence) can simplify many reasoning tasks. Here, we give a short survey of current state.

So far, the most advanced results concerning properties of CI were achieved in the probabilistic framework. The concept of stochastic CI has been studied in probability theory and modern statistics for many years (see [2, 11]). Its importance for the theory of probabilistic ES consists in the fact that any CI–statement can be interpreted

[1]This research was supported by the internal grant n.275105 of the Academy of Sciences of Czech Republic "Conditional independence properties in uncertainty processing".

as certain qualitative relationship among symptoms. This brings the possibility of finding the adequate structural model of such ES. This role of CI was discerned and highlighted by the group around J. Pearl [6] (A.Paz, D.Geiger, T.Verma) but many other researchers dealt more or less explicitly with this concept.

Nevertheless, an analogy of the concept of CI was studied (even earlier) is the theory of relational databases [7]. The counterpart of CI in that theory is the concept of *embedded multivalued dependency* (EMVD). Note that an equivalent concept of *qualitative conditional independence* also appeared in ES theory [8].

Another framework in which the concept of CI appeared, is Spohn's theory of *ordinal conditional functions* [12]. This theory, motivated from philosophical point of view, gives a tool for a mathematical description of the dynamic handling of deterministic epistemology, in this sense it constitutes a counterpart of the probabilistic description of epistemic state. As soon as the concept of CI for ordinal conditional functions was introduced, researchers began to study its properties [4], especially for the special class of natural conditional functions (NCF) called "disbelief function" in [9] or "ranking function" in [3].

One of the most popular approaches to dealing with uncertainty in ES is Dempster–Shafer's theory of belief functions. The concept of CI for variables on which belief functions are defined (i.e. the parallel with concepts of CI studied in probability theory and in the theory of NCFs) was introduced lately by Shenoy [10]. Another definition of (unconditional) independence (also for variables) appeared in [1].

Further framework in which CI can be studied is Zadeh's *possibility theory.* This theory was formulated in the end of seventies as certain model of qualitative description of subjective judgements. Lately, Shenoy introduced the concept of CI also in this field. The above mentioned Shenoy's work [10] gives certain unifying point of view on different calculi dealing with CI. He introduced very abstract concept of *valuation–based system* (VBS) and defined CI for VBSs.

2 Basic Definitions

All above mentioned frameworks for study CI have some common setting. Throughout the paper a collection of nonempty finite sets $\{X_i\,;\ i\in N\}$ is supposed to be given. The index set N is also nonempty and finite. Whenever $\emptyset\neq A\subset N$ the symbol X_A denotes the cartesian product $\prod_{i\in A}X_i$. Power set of a set S will be denoted by $exp\,S$.

The symbol $T(N)$ is reserved for the collection of triplets $\langle A,B,C\rangle$ of pairwise disjoint subsets of N where $A\neq\emptyset\neq B$. Following Pearl [6] we call every subset of $T(N)$ *dependency model over* N. A dependency model is called *semigraphoid* iff it is closed under following four inference rules (called axioms by many authors):

$\langle A,B,C\rangle\ \rightarrow\ \langle B,A,C\rangle$	*symmetry*
$\langle A,B\cup C,D\rangle\ \rightarrow\ \langle A,C,D\rangle$	*decomposition*
$\langle A,B\cup C,D\rangle\ \rightarrow\ \langle A,B,C\cup D\rangle$	*weak union*
$[\,\langle A,C,D\rangle\,\&\,\langle A,B,C\cup D\rangle\,]\ \rightarrow\ \langle A,B\cup C,D\rangle$	*contraction.*

Now, we show how dependency models arise in probability theory, the theory of relational databases and the theory of conditional functions.

DEFINITION 1 (CI–models in probability theory)
Probability measure over N is specified by a nonnegative real function $P : X_N \rightarrow \langle 0, \infty)$ such that $\sum \{P(a); a \in X_N\} = 1$. The formula
$P(A) = \sum \{P(a); a \in A\}$ whenever $A \subset X_N$
defines an additive set function (on $exp\, X_N$) i.e. *probability measure over N*. Whenever $\emptyset \neq S \subsetneq N$ and P is a probability measure over N, then its *marginal measure on* S is a probability measure P^S over S defined as follows:
$P^S(A) = P(A \times X_{N \setminus S})$ for $A \subset X_S$.
Moreover, we put $P^N \equiv P$.
Having a probability measure P over N and a triplet $\langle A, B, C \rangle \in T(N)$ we say that A is conditionally independent of B given C in P and write $A \perp B | C (P)$ iff
$\forall a \in X_A \quad b \in X_B \quad c \in X_C \qquad P^{A \cup B \cup C}(abc) \cdot P^C(c) = P^{A \cup C}(ac) \cdot P^{B \cup C}(bc)$
(take $P^{\emptyset}(\cdot) = 1$). The dependency model $\{\, \langle A, B, C \rangle \in T(N);\, A \perp B | C (P)\,\}$ is then called the *CI–model induced by P*.

The concept of CI in the theory of relational databases is known as embedded multivalued dependency :

DEFINITION 2 (CI–models in the theory of relational databases)
Database relation over N is a nonempty subset of X_N. Having a database relation $\emptyset \neq \mathsf{R} \subset X_N$ and $\emptyset \neq S \subsetneq N$ the *marginal relation* R^S is a database relation over S defined as follows:
$s \in \mathsf{R}^S \Leftrightarrow [\,(s,t) \in \mathsf{R}$ for some $t \in X_{N \setminus S}]$ whenever $s \in X_S$.
Of course, $\mathsf{R}^N \equiv \mathsf{R}$.
Having a database relation R over N and $\langle A, B, C \rangle \in T(N)$ write $A \perp B | C (\mathsf{R})$ iff
$\forall a \in X_A \quad b \in X_B \quad c \in X_C \qquad (a, c) \in \mathsf{R}^{A \cup C} \,\&\, (b, c) \in \mathsf{R}^{B \cup C} \Rightarrow (a, b, c) \in \mathsf{R}^{A \cup B \cup C}$.
The dependency model $\{\langle A, B, C \rangle \in T(N);\, A \perp B | C (\mathsf{R})\}$ is then called the *CI–model induced by* R.

Finally, we define the concept of CI for a special class of ordinal conditional functions, namely so-called natural conditional functions (according to Hunter [4]).

DEFINITION 3 (CI–models in the theory of ordinal conditional functions)
Natural conditional function over N is specified by a nonnegative integer function $\kappa : X_N \rightarrow \{0, 1, 2, \ldots\}$ satisfying $min\, \{\kappa(a); a \in X_N\} = 0$. The formula
$\kappa(A) = min\, \{\kappa(a); a \in A\}$ whenever $\emptyset \neq A \subset X_N$
then defines a set function on $(exp\, X_N) \setminus \{\emptyset\}$ called the *natural conditional function (NCF) over N*. If moreover $\emptyset \neq S \subsetneq N$, then its *marginal NCF* is an NCF over S defined as follows:
$\kappa^S(A) = \kappa(A \times X_{N \setminus S})$ for $A \subset X_S$.
Moreover, $\kappa^N \equiv \kappa$.
Having an NCF κ over N and a triplet $\langle A, B, C \rangle \in T(N)$ write $A \perp B | C (\kappa)$ iff
$\forall\, a \in X_A \quad b \in X_B \quad c \in X_C \qquad \kappa^{A \cup B \cup C}(abc) + \kappa^C(c) = \kappa^{A \cup C}(ac) + \kappa^{B \cup C}(bc)$
(take $\kappa^{\emptyset}(\cdot) = 0$). The dependency model $\{\, \langle A, B, C \rangle \in T(N);\quad A \perp B | C (\kappa)\}$ is called the *CI–model induced by* κ.

3 Comparison

Having introduced the class of CI–models for some specific framework we may ask which inference rules (or axioms) of the form :
$[\,\langle A_1, B_1, C_1\rangle \;\;\&\;\; \ldots \;\;\&\;\; \langle A_n, B_n, C_n\rangle\,] \to \langle A_{n+1}, B_{n+1}, C_{n+1}\rangle$
are *sound* in this particular framework i.e. whether for every instance $\blacklozenge$ of the framework (i.e. probability measure resp. database relation resp. NCF) it holds :
if $[\,A_1 \perp B_1|C_1(\blacklozenge) \;\;\;\&\;\; \ldots \;\;\&\;\;\; A_n \perp B_n|C_n(\blacklozenge)]$ then $A_{n+1} \perp B_{n+1}|C_{n+1}(\blacklozenge)$.

Moreover, we may ask whether there exists a finite *axiomatic characterization* of CI–models i.e. a finite collection of such inference rules characterizing CI–models (for the particular framework) as dependency models satisfying that finite collection of inference rules. Such characterization (even for subclasses of CI–models) would have great importance for reasoning task within the particular framework.

Thus, the classes of CI–models arising in 3 above mentioned frameworks can be compared : we may ask which inference rules are sound in each of the frameworks, whether they differ and whether there exist finite axiomatic characterizations.

At the first sight (according to basic results) the classes of CI-models are very alike: CI-models from all three areas are semigraphoids (for the probabilistic case see [11], for database relations [7], for NCFs [4]). Nevertheless, the classes are indeed different as the following examples show.

Example 1 The following inference rule
$[\,\langle A, B, C \cup D\rangle \,\&\, \langle C, D, A\rangle \,\&\, \langle C, D, B\rangle \,\&\, \langle A, B, \emptyset\rangle\,] \;\longrightarrow\; \langle C, D, \emptyset\rangle$
is sound in probabilistic framework but not in the framework of relational databases and ordinal conditional functions. The probabilistic soundness can be proved using some tools of information theory, for details see [14]. To show that it is not sound for relational databases it suffices to give an example of a database relation R such that the *antecedents* are satisfied i.e. $A \perp B|C \cup D(\mathsf{R}), \ldots, A \perp B|\emptyset(\mathsf{R})$ but the *consequent* is not valid i.e. $\neg[C \perp D|\emptyset(\mathsf{R})]$. To this end we simply take $X_A = \{a, a'\}$, $X_B = \{b, b'\}$, $X_C = \{c, c'\}$, $X_D = \{d, d'\}$ and define R on $X_A \times X_B \times X_C \times X_D$ by the following table (the bullet in a box $[ab, cd]$ means that $(abcd) \in \mathsf{R}$).

R	ab	ab'	$a'b$	$a'b'$
cd	$\bullet$	$\bullet$	$\bullet$	$\bullet$
cd'	$\bullet$			
$c'd$				$\bullet$
$c'd'$				

Similarly we can refute the soundness of this inference rule for NCFs. Take the same sets $X_A, \ldots, X_D$ and define an NCF κ by the following table (the number in a box $[ab, cd]$ is the value $\kappa(abcd)$):

κ	ab	ab'	$a'b$	$a'b'$
cd	0	1	1	2
cd'	1	1	1	1
$c'd$	0	0	0	0
$c'd'$	2	1	1	0

Note that another example that this inference rule fails in the case of NCFs was already given by Spohn [13].

EXAMPLE 2 The following inference rule
$[\langle A, B, C \cup D\rangle \,\&\, \langle C, D, A\rangle \,\&\, \langle C, D, B\rangle\,] \longrightarrow \langle C, D, A \cup B\rangle$
is sound in the framework of relational databases but fails for probabilistic measures and NCFs. Its soundness is easy to see:
Suppose $(a, b, c) \in \mathsf{R}^{A\cup B\cup C}$ and $(a, b, d) \in \mathsf{R}^{A\cup B\cup D}$. As $(a, c) \in \mathsf{R}^{A\cup C}$ and $(a, d) \in \mathsf{R}^{A\cup D}$ by $C \perp D|A(\mathsf{R})$ derive $(a, c, d) \in \mathsf{R}^{A\cup C\cup D}$. Similarly by $C \perp D|B(\mathsf{R})$ get $(b, c, d) \in \mathsf{R}^{B\cup C\cup D}$ and hence using $A \perp B|C \cup D(\mathsf{R})$ finally get $(a, b, c, d) \in \mathsf{R}^{A\cup B\cup C\cup D}$.
To refute the probabilistic soundness take X_A, X_B, X_C, X_D from Example 1 and define on $X_A \times X_B \times X_C \times X_D$ the probability measure P as follows (the value in the box $[ab, cd]$ is $P(abcd)$):

P	ab	ab'	$a'b$	$a'b'$
cd	0.2	0	0	0
cd'	0.2	0	0	0
$c'd$	0.2	0	0	0
$c'd'$	0.1	0.1	0.1	0.1

The counterexample refuting soundness for NCFs was already given in Example 1.

EXAMPLE 3 The following inference rule
$[\langle A, B, C \cup D\,\rangle \,\&\, \langle A, C, B \cup D\rangle] \longrightarrow \langle A, B \cup C, D\rangle$
well–known as *intersection* [6] is sound in the framework of ordinal conditional functions but fails for probabilistic measures and database relations. Its soundness was shown in [12] or [4]. To refute the probabilistic soundness put $D = \emptyset$, take $X_A = \{a, a'\}$, $X_B = \{b, b'\}$, $X_C = \{c, c'\}$ and define on $X_A \times X_B \times X_C$ a probability measure P as follows:

P	bc	bc'	$b'c$	$b'c'$
a	0.5	0	0	0
a'	0	0	0	0.5

The counterexample for relational databases can be obtained easily :
simply take the support $\mathsf{R} = \{x \in X_N;\ P(x) > 0\}$.

Note that the intersection inference rule holds also for strictly positive probability measures. Also further 3 inference rules sound for NCFs (see [13]) hold for strictly positive probability measures.

On the other hand, in all three above mentioned frameworks the following inference rules are sound ($n \geq 3$):
$[\langle A, B_1, B_2\rangle \,\&\ldots\&\, \langle A, B_j, B_{j+1}\rangle \,\&\ldots\&\, \langle A, B_{n-1}, B_n\rangle \,\&\, \langle A, B_n, B_1\rangle] \longrightarrow \langle A, B_2, B_1\rangle$
Moreover, as every proper subset of the collection of antecedents is a CI–model (in all 3 frameworks) this sequence of inference rules can be used to show that for each $n \geq 3$ and every hypothetic complete system $\mathcal{S}$ of sound inference rules there exists an inference rule in $\mathcal{S}$ with at least n antecedents. These results are proved for the probabilistic framework in [15], for relational databases in [7] and for NCFs in [16].

THEOREM For each of the three above mentioned frameworks there exist no finite complete axiomatic characterization of CI–models.

4 Discussion

Let us mention further frameworks for CI which in some sense comprehend the frameworks above – namely possibility theory and Dempster–Shafer theory. The definitions below originate from Shenoy's work [10] where the concept of CI is defined for arbitrary framework satisfying certain general system of axioms for so–called *valuation–based systems.* The main result of that work says that every CI–model arising in such a framework is a semigraphoid.

Definition 4 (CI–models in possibility theory)
Possibility function over N is specified by a real function $\pi : X_N \to \langle 0,1 \rangle$ such that $max\, \{\pi(a);\, a \in X_N\} = 1$. The formula
$\pi(A) = max\, \{\pi(a);\, a \in A\}$ whenever $\emptyset \neq A \subset X_N$
defines a set function on $(exp\, X_N) \setminus \{\emptyset\}$ called *possibility function over* N. Whenever $\emptyset \neq S \subsetneq N$ its *marginal function on* S is a possibility function over S defined as follows:
$\pi^S(A) = \pi(A \times X_{N \setminus S})$ for $\emptyset \neq A \subset X_S$.
Of course, $\pi^N \equiv \pi$.
Having a possibility function π over N and $\langle A,B,C \rangle \in T(N)$ write $A \perp B | C(\pi)$ iff
$\forall \quad a \in X_A \quad b \in X_B \quad c \in X_C \quad \pi^{A \cup B \cup C}(abc) \cdot \pi^C(c) = \pi^{AC}(ac) \cdot \pi^{BC}(bc)$.
For empty C put $\pi^{\emptyset}(c) = 1$. The dependency model $\{\langle A,B,C \rangle \in T(N);\, A \perp B | C(\pi)\}$ is called the *CI–model induced by* π.

This framework in fact involves frameworks for relational databases and NCFs. Indeed, we can assign the following possibility function to each database relation $\mathsf{R} \subset X_N$:
$$\pi_{\mathsf{R}}(A) = \begin{cases} 1 & \text{in case } A \cap \mathsf{R} \neq \emptyset \\ 0 & \text{otherwise.} \end{cases}$$
It makes no problem to verify that the CI–model induced by R coincides with the CI–model induced by π_{R}. Similarly, to every NCF κ over N we can assign the possibility function π_κ as follows :
$\pi_\kappa(A) = e^{-\kappa(A)}$ whenever $\emptyset \neq A \subset X_N$.
This definition ensures that the CI–model induced by κ is identical with the CI–model induced by π_κ. Thus, the class of possibilistic CI–models includes strictly both database CI–models and CI–models arising in NCF theory (the inclusion is proper by Examples 2 and 3).

Shenoy in [10] defines CI also in the framework of Dempster–Shafer theory. This definition specialized to the unconditional case coincides with the concept of independence for evidence measures from [1].

Definition 5 (CI–models in Dempster–Shafer theory)
Basic probability assignment (BPA) over N is specified by a real function
$\mathbf{m} : exp\, X_N \to \langle 0, \infty)$ such that $\sum \{\mathbf{m}(A);\, A \subset X_N\} = 1$ and $\mathbf{m}(\emptyset) = 0$.
The sets $A \subset X_N$ satisfying $\mathbf{m}(A) > 0$ are called *focal elements*. The formula
$Com_{\mathbf{m}}(A) = \sum \{\, \mathbf{m}(B);\, A \subset B\, \}$ whenever $A \subset X_N$
defines so–called *commonality function* which corresponds uniquely to the BPA (i.e.

there exist an inverse formula). Whenever $\emptyset \neq S \subsetneq N$ and $\mathbf{m}$ is a BPA over N, then its *marginal BPA over* S is defined as follows :
$\mathbf{m}^S(A) = \sum \{ \, \mathbf{m}(\mathsf{B}), \quad \mathsf{B} \subset X_N \quad \mathsf{B}^S = A \}$ for $A \subset X_S$
(the projection B^S was introduced in the second definition). Moreover, $\mathbf{m}^N \equiv \mathbf{m}$.
Having a BPA $\mathbf{m}$ over N and a triplet $\langle A, B, C\rangle \in T(N)$ write $A \perp B|C(\mathbf{m})$ iff
$\forall \, \mathsf{E} \subset X_{A\cup B\cup C} \quad Com_{\mathbf{m}^{A\cup B\cup C}}(\mathsf{E}) \cdot Com_{\mathbf{m}^C}(\mathsf{E}^C) = Com_{\mathbf{m}^{A\cup C}}(\mathsf{E}^{A\cup C}) \cdot Com_{\mathbf{m}^{B\cup C}}(\mathsf{E}^{B\cup C})$
(take $Com_{\mathbf{m}^\emptyset}(\cdot) = 1$). The dependency model $\{ \, \langle A, B, C\rangle \in T(N); \, A \perp B|C(\mathbf{m}) \, \}$ is called the *CI–model induced by* $\mathbf{m}$.

Note that the presented definition of CI is in fact one of the equivalent definitions from [10] (use Lemma 3.1(v) there); the definition can be reformulated in terms of BPAs but it would be too complicated. Only important fact is that in case $A \perp B|C(\mathbf{m})$ the focal elements of $\mathbf{m}^{A\cup B\cup C}$ have the form
$\mathsf{F} * \mathsf{G} = \{(a, b, c) \, ; \, (a, c) \in \mathsf{F} \quad (b, c) \in \mathsf{G}\}$ for $\mathsf{F} \subset X_{A\cup C}$ and $\mathsf{G} \subset X_{B\cup C}$ with $\mathsf{F}^C = \mathsf{G}^C$.
The above defined class of CI–models involves both database and probabilistic CI–models. Every database relation $\mathsf{R} \subset X_N$ can be identified with a BPA :

$$\mathbf{m}_{\mathsf{R}}(A) = \begin{cases} 1 & \text{in case } A = \mathsf{R} \\ 0 & \text{otherwise} \end{cases}$$

in such a way that the corresponding CI–models coincide. Similarly, every probability measure P over N defines a BPA $\mathbf{m}_P$ as follows :

$$\mathbf{m}_P(A) = \begin{cases} P(x) & \text{whenever } A = \{x\} \text{ for } x \in X_N \\ 0 & \text{otherwise.} \end{cases}$$

Of course, the CI–model induced by P is the CI–model induced by $\mathbf{m}_P$. Thus, using Examples 1 and 2 we can derive that the presented class of CI–models for BPAs has probabilistic and database CI–models as proper subsets.

Nevertheless, the presented definition does not seem to us to be suitable in the framework of Dempster–Shafer theory. We have two objection.

Firstly, this CI is not "consistent with marginalization". This means that it may happen that for a couple of BPA's $\mathbf{m}_1$ over $A \cup C$ and $\mathbf{m}_2$ over $B \cup C$ which are consonant (i.e. $\mathbf{m}_1^C = \mathbf{m}_2^C$) there exists no BPA $\mathbf{m}$ over $A \cup B \cup C$ such that $\mathbf{m}^{A\cup C} = \mathbf{m}_1$, $\mathbf{m}^{B\cup C} = \mathbf{m}_2$ and $A \perp B|C(\mathbf{m})$. In all other mentioned frameworks such "conditional product" exists and it is uniquely determined.

Secondly, Dempster–Shafer theory was intended to embed both possibility functions and probability measures (see [5]). Concretely, every possibility function π over N is identified with a BPA $\mathbf{m}_\pi$ whose collection of focal elements is a nest i.e.
$\forall \, A, B \subset X_N$ with $\mathbf{m}_\pi(A), \mathbf{m}_\pi(B) > 0$ either $A \subset B$ or $B \subset A$
by means of the relation $\pi(x) = \sum\{\mathbf{m}_\pi(B); x \in B\}$ (for details see [5]). Therefore the concept of CI for BPAs should generalize CI for possibility functions. Nevertheless, this is not true even in the unconditional case. The reason is clear: if we consider two possibilistic BPA's with two focal elements then their product has as focal elements cartesian products of "marginal" focal elements – but this class is not a nest i.e. the product does not represent possibilistic BPA.

We think that the concept CI in Dempster-Shafer theory should comprehend both probabilistic and possibilistic CI and be "consistent with marginalization". However, so far we don't find an adequate definition of CI within this framework.

References

1. L.M. de Campos, J.F. Huete: Independence concepts in upper and lower probabilities. In: Proceedings of IPMU'92 (inter. conf. on Information Processing and Management of Uncertainty in knowledge–based systems, Mallorca, July 6–10, 1992), 129-132

2. A.P. Dawid: Conditional independence in statistical theory. Journal of the Royal Statistical Society B 41, 1–31 (1979)

3. M. Goldszmidt, J. Pearl: Rank–based systems: a simple approach to belief revision, belief update, and reasoning about evidence and actions. In: Proceedings of the 3rd International Conference on Principles of Knowledge Representation and Reasoning, Cambridge, MA, October 1992, 661-672

4. D. Hunter: Graphoids and natural conditional functions. International Journal of Approximate Reasoning 5, 489–504 (1991)

5. G.J. Klir, T.A.Folger: Fuzzy Sets, Uncertainty and Information. Prentice Hall. Englewood Cliffs, N.J. 1988

6. J. Pearl: Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufman. San Mateo, California 1988

7. Y. Sagiv, S.F. Walecka: Subset dependencies and completeness result for a subclass of embedded multivalued dependencies. Journal of the Association for Computing Machinery 29, 103–117 (1982)

8. G. Shafer, P.P. Shenoy, K. Mellouli: Propagating belief functions in qualitative Markov trees. International Journal of Approximate Reasoning 1, 349–400 (1987)

9. P.P. Shenoy: On Spohn's rule for revision of beliefs. International Journal of Approximate Reasoning 5, 149–181 (1991)

10. P.P. Shenoy: Conditional independence in valuation–based systems. School of Business working paper n. 236, University of Kansas, Lawrence 1992, submitted to Annals of Statistics

11. W. Spohn: Stochastic independence, causal independence and shieldability. Journal of Philosophical Logic 9, 73–99 (1980)

12. W. Spohn: Ordinal conditional functions: a dynamic theory of epistemic states. In: W.L. Harper, B. Skyrms (eds.): Causation in Decision, Belief Change, and Statistics, vol. II . Kluwer Academic Publishers. Dordrecht 1988, 105–134

13. W. Spohn: On the properties of conditional independence. to appear in P. Humphreys (eds.): Patrick Suppes - Mathematical Philosopher. Synthese 1993

14. M. Studený: Multiinformation and the problem of characterization of conditional independence relations. Problems of Control and Information Theory 18, 3–16 (1989)

15. M.Studený: Conditional independence relations have no finite complete characterization. In: Transactions of 11-th Prague Conference on Information Theory, Statistical Decision Functions and Random Processes vol.B . Academia. Prague 1992, 377-396

16. M. Studený: Conditional independence and natural conditional functions. submitted to International Journal of Approximate Reasoning