Milan Studený

Attempts at axiomatic description of conditional independence

Kybernetika, Vol. 25 (1989), No. Suppl, 72--79

Persistent URL: `http://dml.cz/dmlcz/125416`

ATTEMPTS AT AXIOMATIC DESCRIPTION OF CONDITIONAL INDEPENDENCE

MILAN STUDENÝ

This paper concerns the problem of characterization of conditional-independence relations (the abbreviation CIR) which arises in connection with probabilistic expert systems. The first two sections contain some lately found properties (axioms) of CIRs; the rest of the paper is a proposal for a proper generalization of the concept of CIR. Since it is defined by means of the concept of multiinformation it is called the M-relation. Some advantages of the M-relations are discussed. Some questions are still open, we express our hypotheses through the paper. The paper is a survey only, the extended version with proofs will be published later.

1. PROBLEM OF CHARACTERIZATION OF CIRs

Throughout the paper we shall assume the following situation. A finite nonempty collection of finite sets X_i, $i \in N$ with card $X_i \geqq 2$ is given. We shall deal with probability measures on $\prod_{i \in N} X_i$ (or equivalently with collections of random variables ξ_i, $i \in N$ where ξ_i takes values from X_i). Note that we shall often restrict ourselves to strictly positive measures. If $A, B \subset N$ are nonempty and disjoint and P is a measure on $\prod_{i \in N} X_i$, then P^A denotes its marginal on $\prod_{i \in A} X_i$ and $P_{A|B}$ the conditional probability on $\prod_{i \in A} X_i$ with respect to $\prod_{i \in B} X_i$.

For any probability measure P on $\prod_{i \in N} X_i$ and nonempty and disjoint $A, B, C \subset N$ we say that A and B are independent under condition C (the notation is $I^P(A; B \mid C)$) iff $P_{A|B \cup C} = P_{A|C}$ or equivalently $P_{B|A \cup C} = P_{B|C}$. For equivalent definitions see [3]. The definition can be extended for empty sets: $I^P(A; B \mid \emptyset)$ means that $P^{A \cup B}$ is the product of P^A and P^B. If $A = \emptyset$ or $B = \emptyset$, then $I^P(A; B \mid C)$ holds by convention.

Given arbitrary probability measure P on $\prod_{i \in N} X_i$ we have introduced the relation $I^P(\cdot; \cdot \mid \cdot)$. on triplets $[A, B, C]$ where A, B, C are pairwise disjoint subsets of N. This relation is called the *CIR corresponding to P* (conditional-independence relation). If it will not cause confusion we shall use $I(A; B \mid C)$ instead of $I^P(A; B \mid C)$.

An analogous definition was in [3] where five properties (axioms) of CIRs were formulated. Note that all sets A, B, C, D in the following formulations are supposed to be disjoint. The first axiom is the symmetry:

$$I(A; B \mid C) \Leftrightarrow I(B; A \mid C) \tag{A.1}$$

Three other axioms can be integrated into the following one:

$$I(A; B \cup C \mid D) \Leftrightarrow [I(A; B \mid C \cup D) \,\&\, I(A; C \mid D)] \tag{A.2}$$

These two axioms hold without the assumption of strict positivity of the measure unlike the next axiom which holds for strictly positive measures only:

$$\text{(P.1)} \qquad [I(A; B \mid C \cup D) \,\&\, I(A; C \mid B \cup D)] \Rightarrow I(A; B \cup C \mid D)$$

Pearl [3] expressed the completeness conjecture, i.e. he believed that these properties of a ternary relation I imply that I coincides with the CIR corresponding to some strictly positive measure. This conjecture was disproved and a further independent axiom (A.3) was found in [5]:

$$\text{(A.3)} \qquad \begin{cases} [I(A; B \mid C \cup D) \,\&\, I(C; D \mid A) \,\&\, I(C; D \mid B) \,\&\, I(A; B \mid \emptyset)] \Leftrightarrow \\ \Leftrightarrow [I(C; D \mid A \cup B) \,\&\, I(A; B \mid C) \,\&\, I(A; B \mid D) \,\&\, I(C; D \mid \emptyset)] \end{cases}$$

Nevertheless, the problem of characterization of CIRs (i.e. finding of "all" properties of CIRs) seems to be significant in the probabilistic expert system theory. To illustrate it let us mention the intensional expert system INES (cf. [4]). According to this approach, the knowledge base of an expert system is modelled by a multidimensional probability measure, while pieces of partial knowledge obtained from experts are described by means of less-dimensional probability measures which should be marginals of the mentioned multidimensional one. Usually it is impossible to store the multidimensional measure in a memory of a computer. This handicap is removed by the help of so-called DSS-approximations which correspond to making some variables (symptoms) conditionally independent. The choice of the DSS-approximation in INES is made from a certain information-theoretical point of view.

The solution of our problem would enable us possible some improvement. Since the notion of conditional independence (or dependence) is easy to interpret we would be able to determine the proper structure of dependences and independences directly by asking experts. By means of the solution of the problem we would be able to decide whether the statements of various experts are contradictory or whether there exists a probabilistic model (i.e. there exists a CIR having prescribed dependences and independences).

2. FURTHER PROPERTIES OF CIRs

Our further investigation detected another independent property of CIRs that concerns strictly positive measures:

$$\text{(P.2)} \quad \begin{bmatrix} I(A; B \mid C \cup D) \,\&\, I(A; B \mid C) \,\&\, I(A; B \mid D) \,\& \\ \& \, I(C; D \mid A \cup B) \,\&\, I(C; D \mid A) \,\&\, I(C; D \mid B) \end{bmatrix} \Rightarrow [I(A; B \mid \emptyset) \,\&\, I(C; D \mid \emptyset)]$$

Moreover, a special property in a binary case (i.e. card $X_i = 2$) was found: (A, B, C are singletons here)

$$\text{(B.1)} \quad [I(A; B \mid C) \,\&\, I(A; B \mid \emptyset)] \Rightarrow [I(A; B \cup C \mid \emptyset) \;\text{ or }\; I(B; A \cup C \mid \emptyset)]$$

Let us mention two papers indirectly concerning the mentioned problem. It is shown in [2] that the "unconditional" – independence relations (i.e. the CIRs restricted to empty conditions: $I(A; B \mid \emptyset)$) can be characterized by three axioms which are more or less deducible from (A.1) – (A.2). It is shown in [1] that in a case of infinite N it can happen that some CIR is not axiomatizable, i.e. the particular CIR corresponding to a certain probability measure cannot be determined by a finite or recursive system of axioms.

3. WHY TO TRY TO GENERALIZE THE CONCEPT OF CIR?

This section contains motivation remarks only and can be skipped. We shall mention some (subjective) reasons for a proper generalization of the concept of CIR.

The first reason is that we have found too many properties of CIRs. We would like to integrate them, i.e. to show that they are deducible from a more general axiom.

Now, let us consider a situation when the actual probability measure P on $\prod_{i \in N} X_i$ is unknown. If we know that sets A, B are independent under a set C we can consider only such measures P for which $I^P(A; B \mid C)$. Unfortunately, there is no simple analogical way how to express the stochastical independence of three sets A, B, C under condition of the fourth set D. Using probability measure it can be defined by the following symmetric expression:

$$P^{A \cup B \cup C \cup D} = P^{A \cup D} \,.\, P^{B \cup D} \,.\, P^{C \cup D} / P^D \,.\, P^D \,.$$

Certainly, it can be expressed by means of the CIR, as for example:

$$I^P(A \cup B; C \mid D) \;\&\; I^P(A; B \mid D) \,.$$

But, there is no simultaneously symmetric and non-redundant expression by means of CIR. Moreover, always at least two demands must be taken.

4. MULTIINFORMATION

Our proposal for a proper generalization of CIRs is defined by means of the concept of multiinformation; therefore it is called M-relation. Now we give the definition of the concept of multiinformation. Let P be a probability measure on $\prod_{i \in N} X_i$ and $A \subset N$ be nonempty. The *multiinformation* of P^A is its relative entropy with respect to the product of its one-dimensional marginals:

$$M(P^A) = H(P^A, \prod_{i \in A} P^{\{i\}}) \quad \text{where} \quad H(Q, R) = \int \ln (\mathrm{d}Q/\mathrm{d}R)\, \mathrm{d}Q \quad \text{for} \quad Q \ll R \,.$$

Having fixed probability measure P we can consider the multiinformation as a function defined on $\exp N$:

$$I_m[A] = M(P^A) \quad \text{for} \quad A \text{ nonempty} \,, \quad I_m[0] = 0 \,.$$

If we want to emphasize that I_m corresponds to some P we shall use the upper index: I_m^P. It is shown in [5] that each I_m satisfies the following conditions:

(1) if $\operatorname{card} A \leqq 1$, then $I_m[A] = 0$

(2) if $A, B \subset N$, then $I_m[A \cup B] + I_m[A \cap B] \geqq I_m[A] + I_m[B]$

(3) if $A \subset B$, then $I_m[A] \leqq I_m[B]$.

Moreover, for a fixed measure P and disjoint A, B, $C \subset N$ the $I^P(A; B \mid C)$ holds iff I_m^P satisfies:

(4) $$I_m^P[A \cup B \cup C] + I_m^P[C] = I_m^P[A \cup C] + I_m^P[B \cup C]$$

Nevertheless, we don't know whether (1)–(3) characterize all possible multiinformation-functions.

5. M-ODERING

The "domain" of M-relation will be determined by means of *M-ordering*. In this section we shall deal with systems of subsets of N for which this ordering will be introduced. As we shall consider even such situations when one set "iterates" in a system several times, we shall describe these systems by means of mappings from $\exp N$ into $\mathbb{N}$ (denoting the set of nonnegative integers).

A mapping $\mathscr{A}$: $\exp N \to \mathbb{N}$ with values 0 and 1 only corresponds to the following system

$$\{S \in \exp N\,;\ \ \mathscr{A}(S) = 1\}\,.$$

Generally, the value $\mathscr{A}(S)$ has significance of multiplicity of S. Besides the ordinary operation $\mathscr{A} + \mathscr{B}$ we shall use the intersection with a subset of N. Thus, given $\mathscr{A}: \exp N \to \mathbb{N}$, $V \subset N$ we define $(\mathscr{A} \wedge V)$: exp. $N \to \mathbb{N}$ as follows

$$(\mathscr{A} \wedge V)(S) = \sum_{T; S = T \cap V} \mathscr{A}(T) \quad \text{if} \quad S \subset V, \quad (\mathscr{A} \wedge V)(S) = 0 \quad \text{otherwise}\,.$$

Finally, for $S \subset N$ we define δ_S: $\exp N \to \{0, 1\}$ by $\delta_S(S) = 1$ and $\delta_S(T) = 0$ whenever $T \neq S$.

Let us denote by Φ the class of all probability measures on $\prod_{i \in N} X_i$. Supposing $\mathscr{A}, \mathscr{B}$: $\exp N \to \mathbb{N}$ we write $\mathscr{A} \prec \mathscr{B}$ iff for all $P \in \Phi$ it holds

$$\sum_{S \subset N} \mathscr{A}(S) \,.\, I_m^P[S] \leqq \sum_{S \subset N} \mathscr{B}(S) \,.\, I_m^P[S]\,.$$

The introduced relation $\prec$ is an ordering with respect to the equivalence:

$$\mathscr{A} \approx \mathscr{B} \quad \text{iff} \quad \mathscr{A}(S) = \mathscr{B}(S) \quad \text{whenever} \quad \operatorname{card} S \geqq 2\,.$$

Thus, we can restrict ourselves to mappings $\mathscr{A}$: $\mathscr{U} \to \mathbb{N}$ where $\mathscr{U} = \{S \subset N;$ $\operatorname{card} S \geqq 2\}$. Let us mention some basic properties:

(5) $$\mathscr{A} \prec \mathscr{B} \quad \text{and} \quad \mathscr{C} \prec \mathscr{D} \quad \text{implies} \quad \mathscr{A} + \mathscr{C} \prec \mathscr{B} + \mathscr{D}$$

(6) $$\mathscr{A} + \mathscr{C} \prec \mathscr{B} + \mathscr{C} \quad \text{implies} \quad \mathscr{A} \prec \mathscr{B}$$

(7) $$\mathscr{A} \leqq \mathscr{B} \quad \text{implies} \quad \mathscr{A} \prec \mathscr{B}; \quad \text{especially} \quad 0 \prec \mathscr{C} \quad \text{for all} \quad \mathscr{C}$$

(8) $$\mathscr{A} \prec \mathscr{B}, \quad V \subset N \quad \text{implies} \quad (\mathscr{A} \wedge V) \prec (\mathscr{B} \wedge V)$$

Example 1. Taking $N = \{a, b, c, d\}$, $\mathscr{A} = \delta_{\{a,c,d\}} + \delta_{\{b,c,d\}}$, $\mathscr{B} = \delta_{\{a,b,c,d\}} + \delta_{\{c,d\}}$ it follows from (2) that $\mathscr{A} \prec \mathscr{B}$.

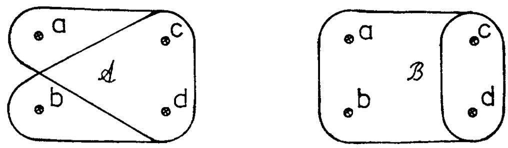

Fig. 1.

For special purposes we can consider a smaller class instead of Φ. For example we can take a class Φ_1 of all probability measures having fixed strictly positive one-dimensional marginals. Generally, such ordering can depend on the choice of the class. Nevertheless, we have proved that the ordering for the class of all strictly positive measures is the same as for Φ. There is a conjecture that having fixed X_i, $i \in N$, the ordering for Φ is the same as for Φ_1.

6. STRONG M-ORDERING

In case of $\mathscr{A}$ such that $\sum_{S \in \mathscr{U}} \mathscr{A}(S) \leqq 2$ the relation $\mathscr{A} \prec \mathscr{B}$ can be characterized as follows. If $\mathscr{A} = \delta_S$, $S \in \mathscr{U}$, then $\mathscr{A} \prec \mathscr{B}$ iff $\mathscr{B}(S') \geqq 1$ for some $S' \supset S$. If $\mathscr{A} = \delta_S + \delta_T$, $S, T \in \mathscr{U}$, $\mathscr{A} \prec \mathscr{B}$, then one of the following three possibilities occurs:

a) $\mathscr{B}(S') \geqq 1$ & $\mathscr{B}(T') \geqq 1$ for some $S' \supset S$, $T' \supset T$ ($\mathscr{B}(S') \geqq 2$ in case $S' = T'$)

b) card $(S \cap T) \leqq 1$ and $\mathscr{B}(U) \geqq 1$ for some $U \supset S \cup T$

c) card $(S \cap T) \geqq 2$ and $\mathscr{B}(U) \geqq 1$ & $\mathscr{B}(V) \geqq 1$ for some $U \supset S \cup T$, $V \supset S \cap T$.

If both $\mathscr{A}$ and $\mathscr{B}$ can be enlarged by some $\mathscr{C}$ in such a way that $\mathscr{A} + \mathscr{C}$ and $\mathscr{B} + \mathscr{C}$ are decomposable into $\mathscr{A}_i \prec \mathscr{B}_i$ where $\sum_{S \in \mathscr{U}} \mathscr{A}_i(S) \leqq 2$, then $\mathscr{A} \prec \mathscr{B}$ according to (5) and (6). More precisely, supposing $\mathscr{A}, \mathscr{B}: \mathscr{U} \to \mathbb{N}$ we write $\mathscr{A} \lhd \mathscr{B}$ iff there exists $\mathscr{A}_i, \mathscr{B}_i, \mathscr{C}: \mathscr{U} \to \mathbb{N}$ $i = 1, \ldots, k$ such that $\mathscr{A} + \mathscr{C} = \sum_i \mathscr{A}_i$. $\mathscr{B} + \mathscr{C} = \sum_i \mathscr{B}_i$, $\mathscr{A}_i \prec \mathscr{B}_i$, $\sum_{S \in \mathscr{U}} \mathscr{A}_i(S) \leqq 2$. We shall call it *strong M-ordering*.

Example 2. Supposing $N = \{a, b, c, d\}$, $\mathscr{A} = \delta_{\{a,c\}} + \delta_{\{b,c\}} + \delta_{\{a,d\}} + \delta_{\{b,d\}}$, $\mathscr{B} = \delta_{\{a,b,c,d\}} + \delta_{\{a,b\}} + \delta_{\{c,d\}}$ it holds $\mathscr{A} \lhd \mathscr{B}$. Indeed, we take $\mathscr{C} = \delta_{\{a,c,d\}} + \delta_{\{b,c,d\}}$ and $\mathscr{A}_1 = \mathscr{C}$, $\mathscr{B}_1 = \delta_{\{a,b\ c\ d\}} + \delta_{\{c\ d\}}$, $\mathscr{A}_2 = \delta_{\{a\ c\}} + \delta_{\{a\ d\}}$, $\mathscr{B}_2 = \delta_{\{a\ c\ d\}}$, $\mathscr{A}_3 = \delta_{\{b,c\}} + \delta_{\{b,d\}}$, $\mathscr{B}_3 = \delta_{\{b,c,d\}}$, $\mathscr{A}_4 = 0$, $\mathscr{B}_4 = \delta_{\{a,b\}}$. Note that we could take $\mathscr{C} = \delta_{\{a,b,c\}} + \delta_{\{a,b,d\}}$ alternatively.

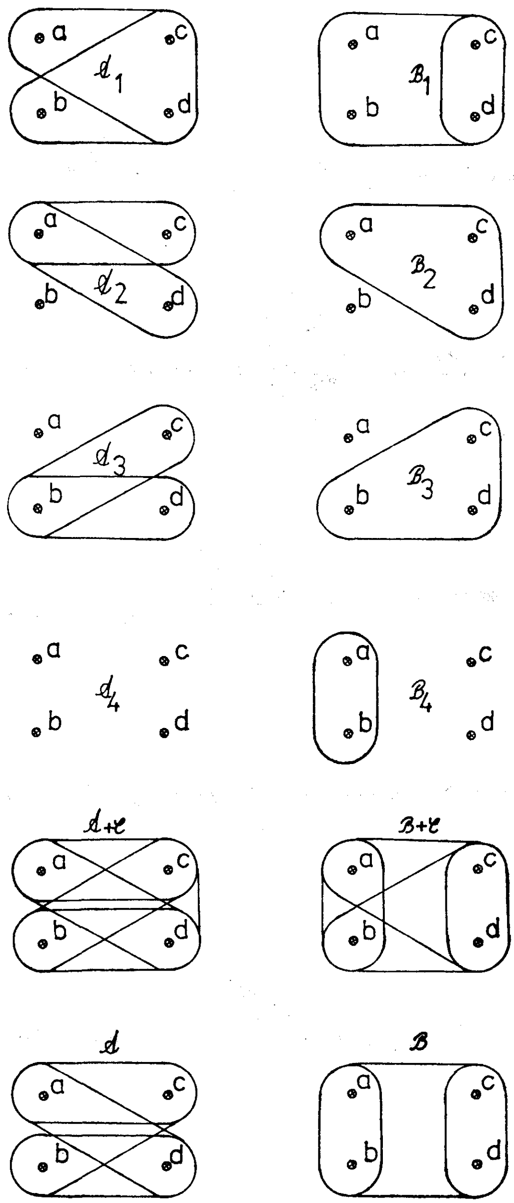

Fig. 2.

Since in case of $\sum_{S\in\mathfrak{U}} \mathscr{A}(S) \leqq 2$ the relation $\mathscr{A} \prec \mathscr{B}$ is described, also the strong M-ordering is characterized. It is really an ordering satisfying (5)−(8). Nevertheless, for the present we don't know whether $\mathscr{A} \prec \mathscr{B}$ implies $\mathscr{A} \lhd \mathscr{B}$, i.e. whether $\prec$ and $\lhd$ coincide.

7. M-RELATION

Let $P \in \Phi$. A binary relation M^P for pairs $\langle \mathscr{A}, \mathscr{B} \rangle$ where $\mathscr{A} \prec \mathscr{B}$ is defined:

$$M^P(\mathscr{A} \mid \mathscr{B}) \quad \text{holds iff} \quad \sum_{S \subset N} \mathscr{A}(S) \,.\, I_m^P[S] = \sum_{S \subset N} \mathscr{B}(S) \,.\, I_m^P[S] \,.$$

We shall called it the *M-relation corresponding to P*.

This relation includes the CIR. Indeed, it follows from (4) that $I^P(A; B \mid C) \Leftrightarrow$ $\Leftrightarrow M^P(\delta_{A \cup C} + \delta_{B \cup C} \mid \delta_{A \cup B \cup C} + \delta_C)$ (see also Example 1). Moreover, the situation treated in the third section can be described as $M^P(\delta_{A \cup D} + \delta_{B \cup D} + \delta_{C \cup D} \mid \delta_{A \cup B \cup C \cup D})$.

8. AXIOMS OF M-RELATIONS

Evidently, for all $P \in \Phi$ (we omit the index P hereafter)

(M.0) if $\mathscr{A} \approx \mathscr{A}'$, $\mathscr{B} \approx \mathscr{B}'$, $\mathscr{A} \prec \mathscr{B}$, then $M(\mathscr{A} \mid \mathscr{B}) \Leftrightarrow M(\mathscr{A}' \mid \mathscr{B}')$

(M.1) $M(\mathscr{A} \mid \mathscr{A})$ holds for all $\mathscr{A}$

(M.2) if $\mathscr{A} \prec \mathscr{B}$, $\mathscr{C} \prec \mathscr{D}$,
then $M(\mathscr{A} + \mathscr{C} \mid \mathscr{B} + \mathscr{D}) \Leftrightarrow [M(\mathscr{A} \mid \mathscr{B}) \,\&\, M(\mathscr{C} \mid \mathscr{D})]$

Moreover, (M.1) and (M.2) imply:

(9) if $\mathscr{A} \prec \mathscr{B}$, then $M(\mathscr{A} \mid \mathscr{B}) \Leftrightarrow M(\mathscr{A} + \mathscr{E} \mid \mathscr{B} + \mathscr{E})$ for all $\mathscr{E}$

(10) if $\mathscr{A} \prec \mathscr{B}$, then $M(\mathscr{A} \mid \mathscr{B} + \mathscr{C}) \Rightarrow M(\mathscr{A} \mid \mathscr{B})$ for all $\mathscr{C}$

(11) if $\mathscr{A} + \mathscr{C} \prec \mathscr{B}$, then $M(\mathscr{A} \mid \mathscr{B}) \Rightarrow M(\mathscr{A} + \mathscr{C} \mid \mathscr{B})$

(12) if $\mathscr{A} \prec \mathscr{B}, \mathscr{B} \prec \mathscr{C}$, then $M(\mathscr{A} \mid \mathscr{C}) \Leftrightarrow [M(\mathscr{A} \mid \mathscr{B}) \,\&\, M(\mathscr{B} \mid \mathscr{C})]$

(13) if $\mathscr{A} \prec \mathscr{B}$, $V \subset N$ such that $\sum_{S, i \in S} \mathscr{A}(S) \leqq 1$ and $\sum_{S, i \in S} \mathscr{B}(S) \leqq 1$
for all $i \in N \setminus V$, then $M(\mathscr{A} \mid \mathscr{B}) \Rightarrow M(\mathscr{A} \wedge V \mid \mathscr{B} \wedge V)$

Example 3. Let $N = \{a, b, c, d\}$ and let us suppose that we got information about a CIR (from experts) in form of the following four facts: $I(\{a\}; \{b\} \mid \{c, d\})$, $I(\{c\}; \{d\} \mid \{a\})$, $I(\{c\}; \{d\} \mid \{b\})$, $I(\{a\}; \{b\} \mid \emptyset)$. We can express them by means of M-relation (as in Example 1) and show using (M.2) and (9) (see Example 2) that this four facts are equivalent to the single fact: $M(\delta_{\{a,c\}} + \delta_{\{a,d\}} + \delta_{\{b,c\}} +$ $+ \delta_{\{b,d\}} \mid \delta_{\{a,b,c,d\}} + \delta_{\{a,b\}} + \delta_{\{c,d\}})$. We see that the M-relation allows us to joint information about structure of a probability measure.

So, (M.2) and (9) enable us to describe the whole M-relation by a single fact $M(\cdot \mid \cdot)$, Another advantage of the M-relation: the well-known concept of decomposable model (see [3]) can be viewed as its special case (we omit details). Moreover, the axioms (A.1)–(A.3) of CIRs can be derived from (M.1) and (M.2). Analogously, the properties (P.1) – (P. 2) for strictly positive measures can be derived from a special property (MP) of M-relations corresponding to strictly positive measures.

But it is not so elegant as (M.2); we regard it as provisional.

Let $\mathscr{A} \lhd \mathscr{B}$, $\mathscr{C} \lhd \mathscr{D}$, $\sum_{S, i \in S} \mathscr{A}(S) = \sum_{S, i \in S} \mathscr{B}(S)$ and $\sum_{S, i \in S} \mathscr{C}(S) = \sum_{S, i \in S} \mathscr{D}(S)$ for all $i \in N$.

Let us denote $\mathscr{A}^{\#} = \mathscr{A} - \min(\mathscr{A}, \mathscr{C})$, $\mathscr{B}^{\#} = \mathscr{B} - \min(\mathscr{B}, \mathscr{D})$,

$\mathscr{C}^{\#} = \mathscr{C} - \min(\mathscr{A}, \mathscr{C})$, $\mathscr{D}^{\#} = \mathscr{D} - \min(\mathscr{B}, \mathscr{D})$.

(MP) If $V \subset N$ satisfies $((\mathscr{A}^{\#} + \mathscr{D}^{\#}) \wedge V) \prec ((\mathscr{B}^{\#} + \mathscr{C}^{\#}) \wedge V)$,

$\sum_{S \in \mathscr{U}} ((\mathscr{A}^{\#} + \mathscr{D}^{\#}) \wedge V)(S) \leqq 2$ and

$\sum_{S, i \in S} (\mathscr{B}^{\#} + \mathscr{C}^{\#})(S) \leqq 1$ $\sum_{S, i \in S} (\mathscr{A}^{\#} + \mathscr{D}^{\#})(S) \leqq 1$ for each $i \in N \setminus V$,

then:

$$[M(\mathscr{A} \mid \mathscr{B}) \& M(\mathscr{C} \mid \mathscr{D})] \Rightarrow M((\mathscr{A}^{\#} + \mathscr{D}^{\#}) \wedge V \mid (\mathscr{B}^{\#} + \mathscr{C}^{\#}) \wedge V).$$

It can be shown by means of the following statement. Supposing $\mathscr{A} \lhd \mathscr{B}$ and $\sum_{S, i \in S} \mathscr{A}(S) = \sum_{S, i \in S} \mathscr{B}(S)$ for all $i \in N$ the fact $M^P(\mathscr{A} \mid \mathscr{B})$ implies $\prod_{S \subset N} P^S(x_S)^{\mathscr{A}(S)} = \prod_{S \subset N} P^S(x_S)^{\mathscr{B}(S)}$. Nevertheless, we don't know whether this statement can be conversed.

REFERENCES

[1] I. Kramosil: A note on nonaxiomatizability of independence relations generated by certain probabilistic structure. Kybernetika *24* (1988), 6, 439–446.

[2] F. Matúš: Independence and Radon projections on compact groups. (in Czech). Ph. D. Thesis, Prague 1988. To appear.

[3] J. Pearl: Markov and Bayes Networks: a Comparison of Two Graphical Representations of Probabilistic Knowledge. Technical Report CSD 860024, R-46-1, University of California, Los Angeles 1986.

[4] A. Perez and R. Jiroušek: Constructing an intensional expert system INES. In: Medical Decision-making: Diagnostic Strategies and Expert Systems (J. V. van Bemmel, F. Gremy and J. Zvárová, eds.), North-Holland, Amsterdam 1985, pp. 307–315.

[5] M. Studený: Multiinformation and the problem of characterization of conditional-independence relations. Problems Control Inform. Theory *18* (1989), 1, 3–16.

RNDr. Milan Studený, CSc., Ústav teorie informace a automatizace ČSAV (Institute of Information Theory and Automation — Czechoslovak Academy of Sciences), Pod vodárenskou věží 4, 182 08 Praha 8. Czechoslovakia.