CI–MODELS ARISING AMONG 4 RANDOM VARIABLES*

Milan STUDENÝ and Pavel BOČEK

ABSTRACT

Let $\xi_1, \xi_2, \xi_3, \xi_4$ be a system of 4 finitely-valued random variables. By the *CI-model* (CI = conditional independence) *induced by* $\{\xi_1, \ldots, \xi_4\}$ we understand the list of triplets of disjoint subsets of $\{1,2,3,4\}$ $\{\langle A, B|C\rangle;\ [\xi_i]_{i\in A}$ is cond. independent of $[\xi_i]_{i\in B}$ given $[\xi_i]_{i\in C}\}$. The subject of this contribution is the problem which lists of such triplets are CI-models induced by a system of 4 finitely-valued random variables.

1. INTRODUCTION

The concept of CI has been studied in probability theory and statistics for many years (Dawid, 1979), (Spohn, 1980), (Mouchart and Rolin, 1984). Its role in probabilistic reasoning was discerned and highlighted by the group around J. Pearl (1988) (A. Paz, D. Geiger, T. Verma), but many other researchers dealt more or less explicitly with this concept (Lauritzen *et al.*, 1990), (Shachter, 1990), (Smith, 1989), (Malvestuto, 1994). Moreover, the concept of CI has appeared to be important also for nonprobabilistic approaches to reasoning (Shenoy, 1994).

The idea of Pearl and Paz (1987) to describe CI-models as 'dependency models' closed under 'inference rules' motivated our research in this area (Studený, 1992), (Matúš, 1992). One of our special goals is to decide which dependency models over $\{1,2,3,4\}$ are CI-models. We hope that the solution of this problem will help us to obtain a

*Supported by the internal grants of Academy of Sciences of Czech Republic n. 27564 and n. 275105.

good view on the general case. Indeed, 3 variables are too little to reveal general aspects of CI and 5 variables are too much to be handled. From this point of view the case of 4 variables is a reasonable compromise – it is sufficiently complicated to show deeper properties of CI but still (although hardly) managable by humans.

The aim of this contribution is to describe the history of that problem, review clearly latest results from (Matúš and Studený, 1995), (Matúš, 1995) and formulate the questions which are still open.

2. BASIC CONCEPTS

Throughout the paper we will deal with 4 random variables $\xi_1, \ldots, \xi_4$ where ξ_i takes values in a finite nonempty set X_i (we will call X_i the *range* of ξ_i). The joint distribution of the collection $\{\xi_1, \xi_2, \xi_3, \xi_4\}$ is a probability distribution over $\{1, 2, 3, 4\}$:

DEFINITION 1 (probability measure over N, marginal measure)
Let N be a *basic index set*. In this paper we have $N = \{1, 2, 3, 4\}$. A *probability measure over* N is specified by a collection of nonempty finite sets $\{X_i, i \in N\}$ and by a probability measure on the cartesian product $\prod_{i\in N} X_i$.
Whenever $\emptyset \neq S \subsetneq N$ define a probability measure over S called the *marginal measure of* P and denoted by P^S as follows:
$P^S(A) = P(A \times \prod_{i\in N\setminus S} X_i)$ whenever $A \subset \prod_{i\in S} X_i$.
Moreover, P^N is defined as P itself.

The focus of our interest is the concept of CI.

DEFINITION 2 (conditional independence)
Suppose that P is the joint distribution of a collection of finite-valued random variables $[\xi_i]_{i\in N}$ (on $\prod_{i\in N} X_i$ concretely). Having a triplet $\langle A, B|C\rangle$ of pairwise disjoint subsets of N, where A, B are nonempty, we say that $[\xi_i]_{i\in A}$ is *conditionally independent of* $[\xi_i]_{i\in B}$ *given* $[\xi_i]_{i\in C}$ iff $\forall\ [x_i]_{i\in N} \in \prod_{i\in N} X_i$
$P^{ABC}([x_i]_{i\in ABC}) \cdot P^C([x_i]_{i\in C}) = P^{AC}([x_i]_{i\in AC}) \cdot P^{BC}([x_i]_{i\in BC})$.[1]
We will also say that the CI-statement $A \perp B|C$ is *valid for* P and write $A \perp B|C(P)$.

[1]The juxtaposition AB is used to denote the union $A \cup B$, $P(a)$ is written instead of $P(\{a\})$ and the convention $P^{\emptyset}([x_i]_{i\in\emptyset}) = 1$ is accepted here.

Remarks
a) This definition involves also the ordinal stochastic independence as a special case when there is no conditioning variable i.e. C is empty.
b) Note that the order of components in the triplet used in this paper differs from (Pearl, 1988) where the conditioning area is placed on the second position. We follow the original notation in probability theory: the conditioning area is on the last position after the separator $|$.

Pearl (1988) proposed to describe CI-structures among random variables by means of the concept of dependency model.

DEFINITION 3 (dependency model, CI-model)
Let us denote by $T(N)$ the set of all ordered triplets $\langle A, B|C\rangle$ where $A, B, C \subseteq N$ are pairwise disjoint and A, B nonempty. Every subset of $T(N)$ will be called a *dependency model over* N.
Let P be a probability measure over N. A dependency model I is called the *model of the* CI-*structure of* P iff I is the set of triplets representing valid CI-statements for P.

The following lemma from (Geiger and Pearl, 1990) or (Studený, 1992) will appear to be very useful in the sequel.

LEMMA 1
The intersection of two CI-models is also a CI-model.[2]

The lemma indicates one of possible ways to description of CI-structures arising among fixed number of variables. One need not to keep the list of all CI-models over N since proper intersections of CI-models can be removed.

DEFINITION 4 (irreducible CI-model)
We will say that a CI-model I is *irreducible* iff it cannot be written as the intersection of two CI-models different from I.

As each CI-model can be written as intersection of irreducible CI-models it suffices focus only on irreducible CI-models. However, one can find even more economic way by considering permutations of the basic index set.

[2]The construction of the corresponding probability distribution used to prove this lemma has first of all theoretical value as it enlarges exponentially ranges of random variables.

Definition 5 (permutation of dependency models)
Let $\pi : N \rightarrow N$ be a permutation of the basic index set N. It can be considered as an autobijection of $T(N)$: one can assign the triplet $\langle \pi(A), \pi(B) | \pi(C) \rangle$ to every $\langle A, B|C \rangle \in T(N)$.[3]
Thus, having dependency models I, J over N we will say that I and J are *permutably equivalent* iff there exists a permutation π of N such that $I = \{ \langle \pi(A), \pi(B) | \pi(C) \rangle; \langle A, B|C \rangle \in J \}$, i.e. I is the image of J by the corresponding autobijection of $T(N)$.

Of course, it is indeed an equivalence relation on the class of dependency models over N which moreover respects CI-models:

Lemma 2
Supposing I and J are permutably equivalent dependency models I is a CI-model iff J is a CI-model.[4]

It makes no problem to derive from the previous lemma that also irreducibility of CI-models is saved by the considered equivalence. Therefore, one need not keep the whole list of irreducible CI-models, it suffices to have only one representative for each equivalence class.

3. INFERENCE RULES

The characterization of CI-models by means of the list of irreducible representatives is one of the possible ways. The dual approach is to describe CI-models as dependency models closed under inference rules.

Definition 6 (inference rule)
In general, an *inference rule* with r antecedents ($r \geq 1$) in an $(r+1)$-nary relation on $T(N)$. But in practice we are interested in inference rules expressed by an informal schema, which defines such an $(r+1)$-nary relation for each basic index set N, like the following examples
$\langle A, B|C \rangle \rightarrow \langle B, A|C \rangle$ symmetry
$\langle A, BC|D \rangle \rightarrow \langle A, C|D \rangle$ decomposition
$\langle A, BC|D \rangle \rightarrow \langle A, B|CD \rangle$ weak union
$[\, \langle A, B|CD \rangle, \langle A, C|D \rangle \,] \rightarrow \langle A, BC|D \rangle$ contraction.

[3] Here $\pi(A) = \{ \pi(x); x \in A \}$ denotes the image of a set A.
[4] Hint: if J is the CI-model induced by $[\xi_i]_{i \in N}$ and π is the pertinent permutation of N, then I is induced by $[\xi_{\pi_{-1}(i)}]_{i \in N}$.

An example is the most effective way of indicating what is meant, but the formal definition is more awkward. So-called *regular inference rule* with r antecedents will be specified by the following items:

a) A finite set of symbols $\mathcal{S} = \{A_1, \dots, A_n\}$, possibly including a special symbol $\emptyset$, is given ($n \geq 3$).

b) For each $k = 1, \dots, r+1$ an ordered triplet $[\mathcal{S}_1^k, \mathcal{S}_2^k, \mathcal{S}_3^k]$ of nonempty disjoint subsets of $\mathcal{S}$ is given. The only requirement is that whenever $\mathcal{S}_i^k$ contains the symbol $\emptyset$ then no other symbol is in $\mathcal{S}_i^k$. The syntactic record of the described inference rule is
$[\,\langle \mathcal{S}_1^1, \mathcal{S}_2^1 | \mathcal{S}_3^1 \rangle, \dots, \langle \mathcal{S}_1^r, \mathcal{S}_2^r | \mathcal{S}_3^r \rangle\,] \;\rightarrow\; \langle \mathcal{S}_1^{r+1}, \mathcal{S}_2^{r+1} | \mathcal{S}_3^{r+1} \rangle$
where each set $\mathcal{S}_i^k$ is depicted by the juxtaposition of its elements. To avoid redundancy we suppose that each symbol in $\mathcal{S}$ is contained in some set $\mathcal{S}_i^k$ and that no couple of different symbols is contained in exactly same collection of sets $\mathcal{S}_i^k$.

c) For every basic set N a set of $(r+1)$-tuples of triplets from $T(N)$ is specified as follows: whenever there exists a substitution mapping m assigning to every symbol $A \in \mathcal{S}$ a subset $m(A) \subset N$ (the empty set is allowed, we even require that $m(\emptyset)$ is the empty set) such that $\{\, m(A_i);\, i = 1, \dots, n\,\}$ is a disjoint collection and for each $k = 1, \dots, r+1$ both $\bigcup\{m(A); A \in \mathcal{S}_1^k\} \neq \emptyset$ and $\bigcup\{m(B); B \in \mathcal{S}_2^k\} \neq \emptyset$, i.e. the triplets
$\langle\, \bigcup\{m(A); A \in \mathcal{S}_1^k\}\,,\, \bigcup\{m(B); B \in \mathcal{S}_2^k\} \,|\, \bigcup\{m(C); C \in \mathcal{S}_3^k\,\rangle$
form an $(r+1)$-tuple of elements of $T(N)$ called *inference instance*. First r triplets (for $k \leq r$) in such an inference instance are called *antecedents*, the last one (corresponding to $k = r+1$) is called the *consequent*. The described regular inference rule is then the set of all its inference instances.

Now, we are to explain how inference rules can be used to describe CI-models.

Definition 7 (probabilistically sound inference rule, semigraphoid) Having a regular inference rule and a dependency model $I \subset T(N)$ we will say that I is *closed under* that inference rule if the consequent (of each inference instance of that rule) belongs to I provided that the antecedents (of that instance) belong to I. An inference rule is *probabilistically sound* iff every CI-model is closed under it.

A dependency model closed under symmetry, decomposition, weak union and contraction mentioned in Definition 6 is called a *semigraphoid*.

Note that all semigraphoid inference rules are probabilistically sound according to well-known basic properties of CI (Dawid, 1979). The original idea of Pearl and Paz (1987) was to describe CI-structures as dependency models closed under a finite number of regular inference rules, concretely Pearl (1988) conjectured that CI-models coincide with semigraphoids.

4. OUR FIRST ATTEMPTS

We started our research in this area by the above mentioned Pearl's conjecture, which was identical with an unpublished Spohn's conjecture from 1980 – as we learned from (Spohn, 1994). Unfortunately, we found in (Studený, 1989) that there exist probabilistically sound inference rules not derivable from the semigraphoid inference rules.[5]

Dealing with the natural question whether CI-models could be characterized as dependency models closed under a finite number of inference rules one can limit attention to inference rules with nonredundant antecedents:

DEFINITION 8 (minimal inference rule)
Consider a probabilistically sound (regular) inference rule with r antecedents. In case it has at least one inference instance $[t_1, \ldots, t_{r+1}] \in T(N)^{r+1}$ such that no proper subset of the antecedent set $\{t_1, \ldots, t_r\}$ has the property that each CI-model containing it contains also the consequent t_{r+1}, we will say that the inference rule is *minimal*, the inference instance will be called minimal, too.

Later we found that even stronger limitation is suitable.

DEFINITION 9 (perfect inference rule)
Consider a probabilistically sound (regular) inference rule with r antecedents. In case it has at least one inference instance $[t_1, \ldots, t_{r+1}] \in T(N)^{r+1}$ where each proper subset of the antecedent set $\{t_1, \ldots, t_r\}$ is a CI-model[6] the inference rule will be called *perfect*.

[5] They are denoted (A.3) in the sequel.

[6] Equivalently: each subset of $\{t_1, \ldots, t_r\}$ of cardinality $r - 1$ is a CI-model.

Of course, a perfect inference rule is necessarily minimal, but the converse is not true. For example, the inference rule
$[\,\langle A,B|E\rangle,\ \langle A,C|BE\rangle\ \langle A,D|CE\rangle\,] \rightarrow \langle A,D|E\rangle$
is a minimal probabilistically sound inference rule[7] which is not perfect since $\{\,\langle\{1\},\{2\}|\emptyset\rangle,\ \langle\{1\},\{3\}|\{2\}\rangle\,\}$ is not a CI-model[8].

As a perfect inference instance cannot be derived using a nonperfect inference rule, all perfect inference rules have to be contained in every complete system of probabilistically sound inference rules. This was the clue which helped in (Studený, 1992) to show that CI-models cannot be characterized by means of a finite number of inference rules. Nevertheless, we showed that CI-models can be characterized by means of a countable number of minimal probabilistically sound inference rules.[9] We hope that nonperfect inference rules can be removed from such a system and dare to formulate:

CONJECTURE
CI-models can be characterized as dependency models closed under (a countable number of) perfect inference rules.

The result on (finite) nonaxiomatizability of CI-models led to an alternative approach to description of CI-models by means of so-called imsets (Studený, 1994, 1995). This approach makes it possible to derive simply further probabilistically sound inference rules. For the case of 4 variables we found the following perfect inference rules.

(A.1) $\langle A,B|C\rangle \longleftrightarrow \langle B,A|C\rangle$

(A.2) $[\,\langle A,B|CD\rangle,\ \langle A,C|D\rangle\,] \longleftrightarrow \langle A,BC|D\rangle$

(A.3) $[\,\langle A,B|CDE\rangle,\ \langle C,D|AE\rangle,\ \langle C,D|BE\rangle,\ \langle A,B|E\rangle\,] \longleftrightarrow$
$\longleftrightarrow [\,\langle C,D|ABE\rangle,\ \langle A,B|CE\rangle,\ \langle A,B|DE\rangle,\ \langle C,D|E\rangle\,]$

(A.4) $[\,\langle A,B|CDE\rangle,\ \langle A,D|BE\rangle,\ \langle C,D|AE\rangle,\ \langle B,C|E\rangle\,] \longleftrightarrow$
$\longleftrightarrow [\,\langle A,D|BCE\rangle,\ \langle A,B|DE\rangle,\ \langle B,C|AE\rangle,\ \langle D,C|E\rangle\,]$

(A.5) $[\,\langle A,C|DE\rangle,\ \langle B,D|CE\rangle,\ \langle B,C|AE\rangle,\ \langle A,D|BE\rangle\,] \longleftrightarrow$
$\longleftrightarrow [\,\langle A,D|CE\rangle,\ \langle B,C|DE\rangle,\ \langle B,D|AE\rangle,\ \langle A,C|BE\rangle\,]$

(A.6) $[\,\langle A,B|CE\rangle,\ \langle A,C|DE\rangle,\ \langle A,D|BE\rangle\,] \longleftrightarrow$
$\longleftrightarrow [\,\langle A,C|BE\rangle,\ \langle A,D|CE\rangle,\ \langle A,B|DE\rangle\,]$

[7] It is derived from the semigraphoid inference rules; substitute $A=\{1\}$, $B=\{2\}$, $C=\{3\}$, $D=\emptyset$ to get the corresponding minimal inference instance.

[8] It imlies $\langle\{1\},\{2,3\}|\emptyset\rangle$ by contraction.

[9] This was done under platonic assumption that all CI-models are captured by the creator of that countable system.

$$\text{(A.7)}\quad [\,\langle A,B|CDE\rangle,\,\langle C,D|ABE\rangle,\,\langle A,C|E\rangle,\,\langle B,D|E\rangle\,] \longleftrightarrow$$
$$\longleftrightarrow [\,\langle A,C|BDE\rangle,\,\langle B,D|ACE\rangle,\,\langle A,B|E\rangle,\,\langle C,D|E\rangle\,]$$

Note that we used here a comperessed notation of inference rules. The rules with the sames set of antecedents are collected in one entry having more than one consequent. Also two-way arrows are used to denote alternative inference rules (with antecedents exchanged for consequents). Thus, (A.2) is a compression of decomposition, weak union and contraction.

At that time we hoped that all perfect inference rules could be produced by the new approach. In (Studený, 1994) we specified a class of *structural semigraphoids* which we had formerly conjectured to coincide with CI-models. We tried to verify it in case of 4 variables.

We found all 37 submaximal structural semigraphoids[10] and found that for our purpose it suffices to show that they are CI-models[11]. Considering permutation (see Definition 5) only 10 construction (for 10 representatives of different classes of permutable equivalence) of probability measures were needed. Unfortunately, we succeeded find only 9 construction, the last type of structural semigraphoids appeared not to be a CI-model (Matúš, 1994). We found later that it can be explained by further independent probabilistically sound inference rules announced in (Studený, 1994).

5. LATEST RESULTS

On the other hand, the hypothesis that the 9 construction above give all irreducible CI-models appeared also untrue. Further 4 types of irreducible CI-models (of course not submaximal!) were found – the list of all known irreducible CI-models will be given in the last section. We proceed in (Matúš and Studený, 1995) and (Matúš, 1995) by finding further independent inference rules. Three basic methods of proving their probabilistic soundness can be distinguished. We will not repeat the proofs, they are in (Matúš, 1995).

[10] A structural semigraphoid is *submaximal* iff the only structural semigraphoid containing it properly is the full class $T(N)$.

[11] As every nonmaximal structural semigraphoid is intersection of submaximal structural semigraphoids, it follows from Lemma 1.

The first method utilizes Corollary 3.6 from (Mouchart and Rolin, 1984) formulated for σ-algebras. Note that this property implies well-known inference rule *intersection* (Pearl, 1988) known to be valid for strictly positive measures. But also in general case it gives some weaker conclusion (not expressible in terms of the classical concept of CI) which combined with other CI-statements may imply the required consequent. That is the principle of the proof of the following collection of perfect probabilistically sound inference rules.

(B.1) $[\langle A,B|C\rangle,\ \langle A,B|D\rangle,\ \langle B,C|A\rangle,\ \langle C,D|\emptyset\rangle\,] \rightarrow \langle A,B|\emptyset\rangle$

(B.2) $[\langle A,B|C\rangle,\ \langle A,C|D\rangle,\ \langle B,C|A\rangle,\ \langle B,D|\emptyset\rangle\,] \rightarrow \langle A,B|\emptyset\rangle$

(B.3) $[\langle A,B|C\rangle,\ \langle A,C|D\rangle,\ \langle C,D|A\rangle,\ \langle B,D|\emptyset\rangle\,] \rightarrow \langle A,B|\emptyset\rangle$

(B.4) $[\langle A,B|C\rangle,\ \langle A,C|D\rangle,\ \langle A,D|C\rangle,\ \langle B,D|\emptyset\rangle\,] \rightarrow \langle A,B|\emptyset\rangle$

(B.5) $[\langle A,B|D\rangle,\ \langle A,C|B\rangle,\ \langle B,D|C\rangle,\ \langle A,D|B\rangle\,] \rightarrow \langle A,B|C\rangle$

(B.6) $[\langle A,B|D\rangle,\ \langle A,C|B\rangle,\ \langle B,D|C\rangle,\ \langle B,D|A\rangle\,] \rightarrow \langle A,B|C\rangle$

(B.7) $[\langle A,B|D\rangle,\ \langle A,C|B\rangle,\ \langle B,D|C\rangle,\ \langle B,C|D\rangle\,] \rightarrow \langle A,B|C\rangle$

(B.8) $[\langle A,B|D\rangle,\ \langle A,C|B\rangle,\ \langle B,D|C\rangle,\ \langle C,D|B\rangle\,] \rightarrow \langle A,B|C\rangle$

(B.9) $[\langle A,B|D\rangle,\ \langle A,C|B\rangle,\ \langle B,D|AC\rangle,\ \langle A,D|B\rangle\,] \rightarrow \langle A,B|CD\rangle$

(B.10) $[\langle A,B|D\rangle,\ \langle A,C|B\rangle,\ \langle B,D|AC\rangle,\ \langle B,D|A\rangle\,] \rightarrow \langle A,B|CD\rangle$

The second method is the principle of uniqueness of a factorizable distribution saying that mutually absolute continous probability measures having the same marginals on sets from a system $\mathcal{S}$ and factorizable with respect to $\mathcal{S}$ coincide. Using this principle probabilistic validity of the following perfect inference rules can be shown.

(C.1) $[\langle A,B|C\rangle,\ \langle A,B|D\rangle,\ \langle A,C|D\rangle,\ \langle C,D|AB\rangle\,] \rightarrow \langle A,B|CD\rangle$

(C.2) $[\langle A,C|D\rangle,\ \langle A,B|D\rangle,\ \langle B,C|A\rangle,\ \langle A,D|BC\rangle\,] \rightarrow \langle A,B|CD\rangle$

(C.3) $[\langle A,B|D\rangle,\ \langle A,C|B\rangle,\ \langle A,C|D\rangle,\ \langle B,D|C\rangle\,] \rightarrow \langle A,B|C\rangle$

The third method consists of calculation with heedful cancelation and employing of 'nonstandard' equivalent definitions of CI. In fact, each of the following perfect inference rules has a specific proof. Note, that (D.1) has appeared also in (Spohn, 1994).

(D.1) $[\langle A,B|C\rangle,\ \langle A,B|D\rangle,\ \langle A,B|\emptyset\rangle,\ \langle C,D|AB\rangle\,] \rightarrow \langle A,B|CD\rangle$

(D.2) $[\langle A,B|C\rangle,\ \langle A,B|D\rangle,\ \langle C,D|A\rangle,\ \langle C,D|\emptyset\rangle\,] \rightarrow \langle A,B|\emptyset\rangle$

(D.3) $[\langle A,B|C\rangle,\ \langle A,C|D\rangle,\ \langle B,D|C\rangle,\ \langle B,D|\emptyset\rangle\,] \rightarrow \langle A,B|\emptyset\rangle$

(D.4) $[\,\langle A,B|C\rangle,\ \langle A,C|D\rangle,\ \langle B,D|A\rangle,\ \langle B,D|\emptyset\rangle\,] \rightarrow \langle A,B|\emptyset\rangle$

6. PRESENT STATE

Thus, we can summarize. In the sequel we give the list of all so far known nonmaximal irreducible CI-models. Of course, it suffices to mention only one representative for each class of permutable equivalence. Moreover, more economic way of description of semigraphoids is based on the following concept.

DEFINITION 10 (elementary triplet)
An *elementary triplet* is every triplet $\langle\{a\},\{b\}|K\rangle \in T(N)$ where $K \subset N$, $a,b \in N \setminus K$, $a \neq b$. The class of elementary triplets over N will be denoted by $E(N)$.

It was shown in (Matúš, 1992):

LEMMA 3
Every semigraphoid $I \subset T(N)$ is determined uniquely by its intersection with the set of elementary triplets $E(N)$. It can be reconstructed from $I \cap E(N)$ by means of the following property:

$$\langle A,B|C\rangle \in I \quad \Leftrightarrow \quad \forall a \in A \quad b \in B \quad C \subset K \subset (A \cup B) \setminus \{a,b\} \quad \langle\{a\},\{b\}|K\rangle \in I \cap E(N).$$

Thus, every CI-model can be represented as a subset of $E(N)$ which has 24 elements in our case $N = \{1,2,3,4\}$.

In the following list we omit braces – every component of an elementary triplet is expressed by the juxtaposition of its elements or by the symbol of the empty set.

$\boxed{1.}$ (20 triplets, submaximal, 6 perm. equiv. representatives)
$\langle 1,3|\emptyset\rangle$, $\langle 2,3|\emptyset\rangle$, $\langle 1,4|\emptyset\rangle$, $\langle 2,4|\emptyset\rangle$, $\langle 3,4|\emptyset\rangle$, $\langle 1,3|2\rangle$, $\langle 2,3|1\rangle$, $\langle 1,4|2\rangle$, $\langle 2,4|1\rangle$, $\langle 1,3|4\rangle$, $\langle 1,4|3\rangle$, $\langle 3,4|1\rangle$, $\langle 2,3|4\rangle$, $\langle 2,4|3\rangle$, $\langle 3,4|2\rangle$, $\langle 1,3|24\rangle$, $\langle 2,3|14\rangle$, $\langle 1,4|23\rangle$, $\langle 2,4|13\rangle$, $\langle 3,4|12\rangle$.

$\boxed{2.}$ (18 triplets, submaximal, 4 perm. equiv. representatives)
$\langle 1,4|\emptyset\rangle$, $\langle 2,4|\emptyset\rangle$, $\langle 3,4|\emptyset\rangle$, $\langle 1,2|3\rangle$, $\langle 1,3|2\rangle$, $\langle 2,3|1\rangle$, $\langle 1,4|2\rangle$, $\langle 2,4|1\rangle$, $\langle 1,4|3\rangle$, $\langle 3,4|1\rangle$, $\langle 2,4|3\rangle$, $\langle 3,4|2\rangle$, $\langle 1,2|34\rangle$, $\langle 1,3|24\rangle$, $\langle 2,3|14\rangle$, $\langle 1,4|23\rangle$, $\langle 2,4|13\rangle$, $\langle 3,4|12\rangle$.

$\boxed{3.}$ (18 triplets, submaximal, 1 perm. equiv. representative)
$\langle 1,2|3\rangle$, $\langle 1,3|2\rangle$, $\langle 2,3|1\rangle$, $\langle 1,2|4\rangle$, $\langle 1,4|2\rangle$, $\langle 2,4|1\rangle$, $\langle 1,3|4\rangle$, $\langle 1,4|3\rangle$, $\langle 3,4|1\rangle$, $\langle 2,3|4\rangle$, $\langle 2,4|3\rangle$, $\langle 3,4|2\rangle$, $\langle 1,2|34\rangle$, $\langle 1,3|24\rangle$, $\langle 2,3|14\rangle$, $\langle 1,4|23\rangle$, $\langle 2,4|13\rangle$, $\langle 3,4|12\rangle$.

$\boxed{4.}$ (18 triplets, submaximal, 1 perm. equiv. representative)
$\langle 1,2|\emptyset\rangle$, $\langle 1,3|\emptyset\rangle$, $\langle 2,3|\emptyset\rangle$, $\langle 1,4|\emptyset\rangle$, $\langle 2,4|\emptyset\rangle$, $\langle 3,4|\emptyset\rangle$, $\langle 1,2|3\rangle$, $\langle 1,3|2\rangle$, $\langle 2,3|1\rangle$, $\langle 1,2|4\rangle$, $\langle 1,4|2\rangle$, $\langle 2,4|1\rangle$, $\langle 1,3|4\rangle$, $\langle 1,4|3\rangle$, $\langle 3,4|1\rangle$, $\langle 2,3|4\rangle$, $\langle 2,4|3\rangle$, $\langle 3,4|2\rangle$.

$\boxed{5.}$ (18 triplets, submaximal, 4 perm. equiv. representatives)
$\langle 1,2|\emptyset\rangle$, $\langle 1,3|\emptyset\rangle$, $\langle 2,3|\emptyset\rangle$, $\langle 1,4|\emptyset\rangle$, $\langle 2,4|\emptyset\rangle$, $\langle 3,4|\emptyset\rangle$, $\langle 1,2|4\rangle$, $\langle 1,4|2\rangle$, $\langle 2,4|1\rangle$, $\langle 1,3|4\rangle$, $\langle 1,4|3\rangle$, $\langle 3,4|1\rangle$, $\langle 2,3|4\rangle$, $\langle 2,4|3\rangle$, $\langle 3,4|2\rangle$, $\langle 3,4|12\rangle$, $\langle 2,4|13\rangle$, $\langle 1,4|23\rangle$.

$\boxed{6.}$ (14 triplets, submaximal, 6 perm. equiv. representatives)
$\langle 1,3|\emptyset\rangle$, $\langle 2,3|\emptyset\rangle$, $\langle 1,4|\emptyset\rangle$, $\langle 2,4|\emptyset\rangle$, $\langle 3,4|\emptyset\rangle$, $\langle 1,3|2\rangle$, $\langle 2,3|1\rangle$, $\langle 1,4|2\rangle$, $\langle 2,4|1\rangle$, $\langle 1,2|34\rangle$, $\langle 1,3|24\rangle$, $\langle 2,3|14\rangle$, $\langle 1,4|23\rangle$, $\langle 2,4|13\rangle$.

$\boxed{7.}$ (12 triplets, submaximal, 1 perm. equiv. representative)
$\langle 1,2|\emptyset\rangle$, $\langle 1,3|\emptyset\rangle$, $\langle 2,3|\emptyset\rangle$, $\langle 1,4|\emptyset\rangle$, $\langle 2,4|\emptyset\rangle$, $\langle 3,4|\emptyset\rangle$, $\langle 1,2|34\rangle$, $\langle 1,3|24\rangle$, $\langle 2,3|14\rangle$, $\langle 1,4|23\rangle$, $\langle 2,4|13\rangle$, $\langle 3,4|12\rangle$.

$\boxed{8.}$ (12 triplets, submaximal, 4 perm. equiv. representatives)
$\langle 1,2|\emptyset\rangle$, $\langle 1,3|\emptyset\rangle$, $\langle 2,3|\emptyset\rangle$, $\langle 1,4|\emptyset\rangle$, $\langle 2,4|\emptyset\rangle$, $\langle 3,4|\emptyset\rangle$, $\langle 1,2|3\rangle$, $\langle 1,3|2\rangle$, $\langle 2,3|1\rangle$, $\langle 2,3|14\rangle$, $\langle 1,3|24\rangle$, $\langle 1,2|34\rangle$.

$\boxed{9.}$ (12 triplets, submaximal, 4 perm. equiv. representatives)
$\langle 2,3|\emptyset\rangle$, $\langle 2,4|\emptyset\rangle$, $\langle 3,4|\emptyset\rangle$, $\langle 2,3|1\rangle$, $\langle 2,4|1\rangle$, $\langle 3,4|1\rangle$, $\langle 1,2|34\rangle$, $\langle 1,3|24\rangle$, $\langle 2,3|14\rangle$, $\langle 1,4|23\rangle$, $\langle 2,4|13\rangle$, $\langle 3,4|12\rangle$.

$\boxed{10.}$ (5 triplets, not submaximal, 6 perm. equiv. representatives)
$\langle 1,2|34\rangle$, $\langle 1,2|3\rangle$, $\langle 1,2|4\rangle$, $\langle 3,4|1\rangle$, $\langle 3,4|2\rangle$.

$\boxed{11.}$ (4 triplets, not submaximal, 6 perm. equiv. representatives)
$\langle 1,2|34\rangle$, $\langle 1,2|3\rangle$, $\langle 1,2|4\rangle$, $\langle 3,4|\emptyset\rangle$.

$\boxed{12.}$ (4 triplets, not submaximal, 24 perm. equiv. representatives)
$\langle 1,2|34\rangle$, $\langle 1,2|4\rangle$, $\langle 2,3|1\rangle$, $\langle 3,4|2\rangle$.

$\boxed{13.}$ (4 triplets, not submaximal, 24 perm. equiv. representatives)
$\langle 1,2|34\rangle$, $\langle 1,2|4\rangle$, $\langle 2,3|1\rangle$, $\langle 3,4|\emptyset\rangle$.

The corresponding constructions of probability measures follow. The sets X_i $(i = 1, \ldots, 4)$ are taken either $\{0\}$ or $\{0,1\}$ or $\{0,1,2\}$ or even $\{0,1,2,3\}$ – what option is concretly chosen is clear from the context. The measures are defined on their support only; the order of variables in quadruplets is 1,2,3,4.

$\boxed{1.}$
$(0,0,0,0) \to 1/2$
$(1,1,0,0) \to 1/2$

$\boxed{2.}$
$(0,0,0,0) \to 1/2$
$(1,1,1,0) \to 1/2$

$\boxed{3.}$
$(0,0,0,0) \to 1/2$
$(1,1,1,1) \to 1/2$

$\boxed{4.}$
$(0,0,0,0) \to 1/8$
$(0,0,1,1) \to 1/8$
$(0,1,0,1) \to 1/8$
$(0,1,1,0) \to 1/8$
$(1,0,0,1) \to 1/8$
$(1,0,1,0) \to 1/8$
$(1,1,0,0) \to 1/8$
$(1,1,1,1) \to 1/8$

$\boxed{5.}$
$(0,0,0,0) \to 1/4$
$(0,1,1,0) \to 1/4$
$(1,0,1,0) \to 1/4$
$(1,1,0,0) \to 1/4$

$\boxed{6.}$
$(0,0,0,0) \to 1/4$
$(0,0,1,1) \to 1/4$
$(1,1,0,1) \to 1/4$
$(1,1,1,0) \to 1/4$

$\boxed{7.}$
$(0,0,0,0) \to 1/9$
$(0,1,1,1) \to 1/9$
$(0,2,2,2) \to 1/9$
$(1,0,1,2) \to 1/9$
$(1,1,2,0) \to 1/9$
$(1,2,0,1) \to 1/9$
$(2,0,2,1) \to 1/9$
$(2,1,0,2) \to 1/9$
$(2,2,1,0) \to 1/9$

$\boxed{8.}$
$(0,0,0,0) \to 1/8$
$(0,0,1,1) \to 1/8$
$(0,1,0,2) \to 1/8$
$(0,1,1,3) \to 1/8$
$(1,0,0,3) \to 1/8$
$(1,0,1,2) \to 1/8$
$(1,1,0,1) \to 1/8$
$(1,1,1,0) \to 1/8$

$\boxed{9.}$
$(0,0,0,0) \to 1/4$
$(1,0,1,1) \to 1/4$
$(2,1,0,1) \to 1/4$
$(3,1,1,0) \to 1/4$

$\boxed{10.}$
$(0,0,0,0) \to 1/6$
$(1,0,0,0) \to 1/6$
$(0,0,0,1) \to 1/6$
$(0,0,1,0) \to 1/4$
$(0,0,1,1) \to 1/8$
$(0,1,1,1) \to 1/8$

$\boxed{11.}$
$(0,0,0,0) \to 1/8$
$(1,0,0,0) \to 1/8$
$(0,0,0,1) \to 1/4$
$(0,0,1,0) \to 1/4$
$(0,0,1,1) \to 1/8$
$(0,1,1,1) \to 1/8$

$\boxed{12.}$
$(1,0,0,0) \to 1/8$
$(1,1,0,0) \to 1/8$
$(1,0,1,0) \to 1/8$
$(1,1,1,0) \to 1/4$
$(1,1,0,1) \to 1/8$
$(0,1,1,1) \to 1/4$

$\boxed{13.}$
$(0,0,0,0) \to 1/8$
$(0,1,0,0) \to 1/8$
$(0,1,1,0) \to 1/4$
$(1,0,0,1) \to 1/4$
$(0,0,1,1) \to 1/4$

However, so far we are not sure whether the previous list is complete. Almost all other semigraphoids are excluded by inference rules mentioned in the preceding section, but still we are not sure about the following three cases.

Open Questions

We would like to know whether the following dependency models are CI-models or not:

a) $\langle A, B|C\rangle$, $\langle A, D|B\rangle$, $\langle B, C|AD\rangle$, $\langle A, D|\emptyset\rangle$

b) $\langle A, B|C\rangle$, $\langle A, D|B\rangle$, $\langle B, C|AD\rangle$, $\langle C, D|B\rangle$

c) $\langle A, B|C\rangle$, $\langle A, D|B\rangle$, $\langle B, C|AD\rangle$, $\langle A, D|\emptyset\rangle$, $\langle C, D|B\rangle$

Note that the only CI-statement which could be the consequent of a prospective probabilistically sound inference rule is $\langle A, B|CD\rangle$.

We tried to answer especially the first question and found that the inference rule

$[\,\langle A, B|C\rangle, \langle A, D|B\rangle, \langle B, C|AD\rangle, \langle A, D|\emptyset\rangle\,] \rightarrow \langle A, B|CD\rangle$

is prob. sound in some special situation – when card $\prod_{i\in D} X_i = 2$ and the measure is strictly positive.[12]

As complete list of CI-models contains more than 18 000 items we decided to create a database SGPOKUS of all semigraphoids among 4 variables where it is indicated which semigraphoids is already known to be a CI-model, which one is known not to be a CI-model and for which is this question still open.

Notice This paper is intended only for the workshop WUPES'94 - it has informal character, describing unfinished work. We do not plan to accept the offer of organizators and submit it in this form as publication in a special volume of a journal.

REFERENCES

Dawid, A.P. (1979) Conditional independence in statistical theory, Jounal of the Royal Statistical Society B, 41, 1-31.

Geiger, D., and Pearl, J. (1990) Logical and algorithmic properties of independence and their application to Bayesian networks, Annals of Mathematics and Artificial Intelligence, 2,

[12]The proof is too long and complicated to be presented here.

165-178.

Lauritzen, S.L., Dawid, A.P., Larsen, B.N. and Leimer, H.-G. (1990) Independence properties of directed Markov fields, Networks, 20, 491-505.

Malvestuto, F.M. (1994) Testing implication of loglinear hierarchical models for probability distributions, to appear in Statistics and Computing.

Matúš, F. (1992) Ascending and descending conditional independence relations, in Information Theory, Statistical Decision Functions and Random Processes: transactions of 11th Prague Conference B (S. Kubík and J.Á. Víšek eds.), Kluwer, Dordrecht (also Academia, Prague), 189-200.

Matúš, F. (1994) Probabilistic conditional independence structures and matroid theory: background, International Journal of General Systems, 22, 185-196.

Matúš, F. (1995) Conditional independences among four random variables II., accepted in Combinatorics, Probability and Computing.

Matúš, F. and Studený, M. (1995) Conditional independences among four random variables I., accepted in Combinatorics, Probability and Computing.

Mouchart, M. and Rolin, J.-M. (1984) A note on conditional independence with statistical application, Statistica, 44, 557-584.

Pearl, J. and Paz, A. (1987) Graphoids: a graph-based logic for reasoning about relevance relations, in Advances in Artificial Intelligence II (B. Du Boulay, D. Hogg, and L. Steels eds.), North-Holland, Amsterdam, 357-363.

Pearl, J. (1988) Probabilistic Reasoning in Intelligent Systems, Morgan Kaufman, San Mateo, CA.

Shachter, R.D. (1990) An ordered examination of influence diagrams, Networks, 20, 535-563.

Shenoy, P.P. (1994) Conditional independence in valuation-based systems, International Journal of Approximate Reasoning, 10, 203-234.

Smith, J.Q. (1989) Influence diagrams for statistical modelling, Annals of Statistics, 17, 654-672.

Spohn, W. (1980) Stocastic independence, causal independence and shieldability, Journal of Philosophical Logic, 9, 73-99.

Spohn, W. (1994) On the properties of conditional independence, in Patrick Suppes: Scientific Philosopher (P. Humphreys ed.), Kluwer, Dordrecht, 173-196.

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

Studený, M. (1992) Conditional independence relations have no finite complete characterization, in Information Theory, Statistical Decision Functions and Random Processes: transactions of 11th Prague Conference B (S. Kubík and J.Á. Víšek eds.), Kluwer, Dordrecht (also Academia, Prague), 377-396.

Studený, M. (1994) Structural semigraphoids, International Journal of General Systems, 22, 207-217.

Studený, M. (1995) Description of structures of stochastic conditional independence by means of faces and imsets, a series of 3 papers to appear in International Journal of General Systems.

Address correspondence to
Milan Studený
Department of Decision–Making Theory
Inst. Inform. Theory and Automation
Pod vodárenskou věží 4
182 08 Prague 8, Czech Republic

E-mail: studeny@utia.cas.cz