Description of conditional independence structures by means of imsets: a connection with product formula validity*

Milan Studený
Department of Decision Making Theory, Institute of Information Theory and Automation, Pod vodárenskou věží 4, 182 08 Prague 8, Czech republic.
(studeny@cspgas11.bitnet)

A new approach to mathematical description of structures of stochastic conditional independence, namely by means of so-called imsets, is presented (imset is an abbreviation for **i**nteger-valued **multiset**). It is shown how it is related to the "classical" approach, namely by means of dependency models or semigraphoids. The main result consists in the theorem saying that a probability measure has certain conditional independence structure (CI-structure) if and only if it satisfies the corresponding product formula.

AMS classification : 68T 30 , 62B 10.

INTRODUCTION

The interest in stochastic conditional independence takes its origin from the theory of probabilistic expert systems. To put it shortly, any conditional independence statement can be interpreted as a certain qualitative relationship among symptoms. Therefore there exists a possibility to determine a proper structural model of the probabilistic expert system which is easy to understand. In fact, in the theoretical background of various approaches to qualitative description of probabilistic models (influence diagrams , Markov nets) the concept of *conditional independence* (CI) is hidden. To the best of my knowledge, the importance of CI for probabilistic expert systems was at first highlighted by Pearl [16] but there exist other approaches which more or less explicitly deal with CI [17,18,19,20,26].

The "classical" approach to description of CI-structures (if graphical approaches are omitted) used the concept of dependency model or of semigraphoid [15]. It motivated attempts at "axiomatization" of CI i.e. to characterize relationships among CI-statements in a simple syntactic way. Nevertheless, as proved in [23] there exists no simple dimension-independent deductive system describing relationships among CI-statements. This fact motivated a new approach to description of CI-structures [25], namely by means of imsets. It promises to remove the above mentioned drawbacks. The aim of this article is to give another view on this type of description of CI-stuctures by showing that the complying with such a model of CI-structure is equivalent to the validity of certain product formula. This equivalent definition could make the interpretation of these models more natural.

*This work was made in frame of the project "Explanatory power of probabilistic expert systems: theoretical backgroud" , internal grant of Czechoslovak Academy of Sciences, n.27510.

NOTATION AND BASIC ARRANGEMENTS

Throughout this paper the following situation will be delt with:
A finite set N having at least two elements called the *basic set* is given, i.e. $2 \leq card\, N < \infty$. Having disjoint sets $A, B \subset N$ the juxtaposition AB will stand for their union $A \cup B$ in many examples. The class of all subsets of N will be denoted by $exp\, N$, the class of its *nontrivial subsets* i.e. subsets having at least two elements will be denoted by $\mathcal{U}$: $\quad \mathcal{U} = \{S \subset N,\ card\, S \geq 2\}$.
Having a set $T \subset N$ its *identificator* i.e. the zero-one function on $exp\, N$ (possibly restricted to $\mathcal{U}$) is defined as follows:

$$\delta_T(S) = \begin{cases} 1 & \text{in case } S = T \\ 0 & \text{in case } S \neq T. \end{cases}$$

Having a collection of nonempty finite sets $\{X_i\ ;\ i \in N\}$, an element of the corresponding cartesian product $x \in \prod_{i \in N} X_i$ and a set $\emptyset \neq S \subset N$ the projection of x to $\prod_{i \in S} X_i$ will be denoted by x_S:

$x_S \in \prod_{i \in S} X_i \quad$ is specified by $\quad [\ \forall i \in S \ \ (x_S)_i = x_i\ ]$.

Having a pair a, b of real functions on a finite set X their *scalar product* will be denoted by $\langle a, b \rangle$:

$\langle a, b \rangle = \sum_{x \in X} a(x) \cdot b(x)$.

The set of integers will be denoted by $\mathbb{Z}$, the set of nonnegative integers (including zero) by $\mathbb{Z}^+$, the set of (strictly) positive integer (sometimes called natural numbers) by $\mathbb{N}$ and the set of real numbers by $\mathbb{R}$.
For the sake of brevity, having a probability measure Q on Y (or a function on $exp Y$) and $y \in Y$ the symbol $Q(y)$ will be often used instead of $Q(\{y\})$.
Having a function $w : exp\, N \to \mathbb{Z}$ its positive and negative parts will be denoted by w_+ and w_-:

$w_+(S) = max\, \{w(S), 0\} \qquad$ for $S \subset N$
$w_-(S) = max\, \{-w(S), 0\} \qquad$ for $S \subset N$.

1. THE PRODUCT FORMULA GIVEN BY AN IMSET

This section contains basic definitions and the corresponding comments only. Namely, the concepts of imset, its natural extension and finite-domain probability measure are specified. Finally, the product formula given by an imset is introduced and a simple example given.

Def. 1 (imset)
Every integer-valued function on $\mathcal{U}$ is called *imset* (*on* $\mathcal{U}$).
The class of all imsets will be denoted by $Z(\mathcal{U})$.
Nonnegative imsets will be called *multisets*, their class will be denoted by $Z^+(\mathcal{U})$.
Basic operations with imsets like summing, subtracting and multiplying by integers are defined coordinatewisely. An imset $u \in Z(\mathcal{U})$ is called *normalized* iff the collection of numbers $\{u(S); S \in \mathcal{U}\}$ has no common prime divisor.

Trivial examples of imsets are *zero imset* (a function ascribing zero to each set from $\mathcal{U}$) denoted by 0 and identificators for $T \in \mathcal{U}$.

Remark The term multiset is borrowed from Aigner's book about combinatorial theory [1] while the word imset is our abbreviation for **i**nteger-valued **m**ulti**set**.

In some cases it will be convenient to regard imsets on $\mathcal{U}$ as functions on $exp\,N$. The correctness of the definition of the right extension is based on the following lemma; its proof is left to the reader.

Lemma Every imset $u \in Z(\mathcal{U})$ has uniquely determined extension
$\overline{u} : exp\,N \to \mathbb{Z}$ satisfying the following two conditions:
$(N.1)$ $\sum\{\overline{u}(S); S \subset N\} = 0$
$(N.2)$ $\forall r \in N \quad \sum\{\overline{u}(S);\ r \in S \subset N\} = 0.$
This adjudgement defines a one-to-one correspondence between $Z(\mathcal{U})$ and the class of integer-valued functions on $exp\,N$ satisfying $(N.1) - (N.2)$.

<u>Def. 2</u> (natural extension)
Having an imset u on $\mathcal{U}$ the uniquely determined integer-valued function $\overline{u}$ on $exp\,N$ extending it and satisfying the normalization conditions $(N.1) - (N.2)$ will be called the *natural extension of* u. It will be always denoted by $\overline{u}$ (overline the original symbol).

The focus of our study are CI-structures of finite number of random variables (i.e. random systems). Nevertheless, in this article we limit ourselves to finite-valued random variables. It is a common custom in literature to allude to random variables but in fact deal with probability measures, namely their distributions. In the sequel, we decided both to allude to and deal with probability measures.

Now, the relevant class of probability measures will be specified. As they serve as distributions of random systems indexed by the basic set N their domains are cartesian products indexed by N.

<u>Def. 3</u> (probability measure over N)
A *probability measure over* N (with finite domain) is specified by a collection of <u>nonempty finite</u> sets $\{X_i\ ;\ i \in N\}$ and by a probability measure on the cartesian product $\prod_{i \in N} X_i$.
Whenever $\emptyset \neq S \subsetneq N$ the *marginal measure of* P is the probability measure P^S on $\prod_{i \in S} X_i$ defined by: $P^S(A) = P(A \times \prod_{i \in N \setminus S} X_i)$ whenever $A \subset \prod_{i \in S} X_i$.
It will be always denoted by the symbol of the original measure endowed with the upper index identifying the marginal space. Moreover, the marginal measure P on $\prod_{i \in N} X_i$ is introduced as P itself, i.e. $P^N \equiv P$.

As mentioned above every assumption concerning CI-stucture of a probability measure is equivalent to the validity of certain product formula. Now, we are going to explain how these formulas look.

Def. 4 (product formula given by an imset)
Let P be a probability measure over N and u be an imset on $\mathcal{U}$. Say that P *satisfies the product formula given by* u iff it holds:

$$\forall x \in \prod_{i \in N} X_i \quad \prod_{\emptyset \neq S \subset N} (P^S(x_S))^{\overline{u}_+(S)} = \prod_{\emptyset \neq S \subset N} (P^S(x_S))^{\overline{u}_-(S)} \tag{1.1}$$

The condition (1.1) is the above mentioned formula.

Remark Another possible way how to write product formulas is to introduce the conventional symbol $P^{\emptyset}(x_{\emptyset})$ for 1 and write instead of (1.1) :

$$\forall x \in \prod_{i \in N} X_i \quad \prod_{S \subset N} (P^S(x_S))^{\overline{u}_+(S)} = \prod_{S \subset N} (P^S(x_S))^{\overline{u}_-(S)} \tag{1.2}$$

To illustrate this concept a simple example is given.

Example 1
Consider $S, T \in \mathcal{U}$ with $S \cap T = \emptyset$ and $S \cup T = N$. Put $u = \delta_N - \delta_S - \delta_T$. Then $\overline{u} = u + \delta_{\emptyset}$ and the product formula given by u looks:

$$\forall x \in \prod_{i \in N} X_i \quad P(x) = P^S(x_S) \cdot P^T(x_T) \tag{1.3}$$

2. DESCRIPTION OF CI-STRUCTURES

In this section the concept of CI is recalled and the "classical" ways to description of CI-structures are mentioned, especially by means of dependency models. The reason to develop the new approach to description of CI-structures from [25] is explained.

Def. 5 (conditional independence)
Let P be a probability measure over N and $\langle A, B, C \rangle$ is a triplet of pairwise disjoint subsets of N where A and B are nonempty. Say that A is *conditionally independent of* B *given* C *in* P and write $A \perp B | C\ (P)$ iff

$$\forall x \in \prod_{i \in N} X_i \quad P^{ABC}(x_{ABC}) \cdot P^C(x_C) = P^{AC}(x_{AC}) \cdot P^{BC}(x_{BC}).$$

Another phrase " *P obeys the triplet $\langle A, B, C \rangle$* " will be often used in the sequel.

There are many equivalent formulations of the statement $A \perp B | C\ (P)$, for example:

- $\forall a, \tilde{a} \in \prod_{i \in A} X_i \quad b, \tilde{b} \in \prod_{i \in B} X_i \quad c \in \prod_{i \in C} X_i$
  $P^{ABC}([a,b,c]) \cdot P^{ABC}([\tilde{a},\tilde{b},c]) = P^{ABC}([a,\tilde{b},c]) \cdot P^{ABC}([\tilde{a},b,c])$
- there exist functions $f : \prod_{i \in AC} X_i \to \mathbb{R}$ and $g : \prod_{i \in BC} X_i \to \mathbb{R}$ such that
  $\forall x \in \prod_{i \in N} X_i \quad P^{ABC}(x_{ABC}) = f(x_{AC}) \cdot g(x_{BC})$
- $\forall c \in \prod_{i \in C} X_i$ with $P^C(c) > 0$ the conditional probability $P_{AB|C}(\cdot|c)$ is a product measure on $(\prod_{i \in A} X_i) \times (\prod_{i \in B} X_i)$.

The last condition leads directly to the common interpretation of $A \perp B|C$: "getting know the values of variables from C the variables from A and B become independent i.e. their probabilistic pieces of information become unrelevant". Thus, the information about CI-structure can be obtained from experts too.

To approach CI to human understanding the CI-structures were usually described by means of graphs in literature. Two trends are distinguishable: by means of undirected graphs (this stems from Markov field theory, the corresponding graph is called Markov net [2,6,8,12] and by means of directed acyclic graphs (the long tradition started by geneticist S. Wright [28] led to the concepts of influence diagram [5,17,18,19] and recursive models [7,27]. Nevertheless both graphical approaches cannot describe all possible probabilistic CI-structures.

Thus, another natural way was proposed: to describe a CI-structure simply by the list of valid CI-statements (i.e. triplets obeyed by the corresponding probability measure). This led to the concept of dependency model introduced by Pearl and Paz [15]; their definition is slightly modified here:

Def. 6 (dependency model, model of CI-structure)
a) Denote by $T(N)$ the set of triplets $\langle A, B, C \rangle$ of pairwise disjoint subsets of N where A and B are nonempty. Every subset of $T(N)$ will be called a *dependency model over* N.
b) Let P be a probability measure over N and I a dependency model over N. Say that I is a *submodel of* CI-*structure of* P iff P obeys every triplet from I.
Further, say that I is the *model of* CI-*structure of* P iff I is exactly the set of triplets obeyed by P.

This terminology emphasizes the presented view on dependency models. Note that authors dealing with dependency models have used also various another phrases:
"I is induced by P" in [26], "P is perfect for I" in [4], "I is conditional independence relation corresponding to P" in [22].

Owing to well-known properties of CI (treated by Dawid [3] resp. Spohn [21] resp. Smith [19]) some of dependency models cannot serve as (complete) models of CI-structures. Therefore Pearl and Paz [15] introduced the concept of semigraphoid (= dependency model satisfying the above mentioned properties) to describe CI-structures. As semigraphoids were defined as dependency models closed under 4 inference rules (called axioms by Pearl and Paz) it gives a deductive mechanism to infer valid consequences of input information about CI-structure.

Unfortunately, the original hypothesis from [16] that semigraphoids coincide with the (complete) models of CI-structures appeared untrue [22]. Later, we even found that models of CI-structures cannot be described as dependency models closed under finite number of inference rules [23]. This was strengthened by Geiger and Pearl [4] who showed that "disjunctive" inference rules cannot bring help.

These results led us to an attempt to develop an alternative way to description of CI-structures, namely by means of faces and imsets [25]. The aim of this paper is to give an equivalent view on "imsetal" models of CI-structures which brings an easier interpretation.

3. INFORMATION-THEORETICAL APPROACH

This section recalls the information-theoretical concepts of entropic and multiinformation function and indicates how they enable to describe CI-structures by means of imsets. An initial connection with product formulas is established.

Entropic and multiinformation functions are real functions on $exp\, N$. In fact, the value of entropic function for a set S is the entropy of the marginal measure P^S while the value of multiinformation function is the relative entropy of P^S with respect to the product of its one-dimensional marginals.

Def. 7 (entropic function, multiinformation function)
Let P a probability measure over N.
Its *entropic function* $H : exp\, N \to \mathbb{R}$ is defined as follows:

$$\begin{aligned} H(\emptyset) &= \quad 0 \\ H(S) &= \sum_{x, P(x)>0} P(x) \cdot ln\,(1/P^S(x_S)) \qquad \text{for } \emptyset \neq S \subset N. \end{aligned}$$

Its *multiinformation function* $M : exp\, N \to \mathbb{R}$ is defined by:

$$\begin{aligned} M(\emptyset) &= \quad 0 \\ M(S) &= \sum_{x, P(x)>0} P(x) \cdot ln\,(P^S(x_S)/ \prod_{i \in S} P^{\{i\}}(x_i)) \qquad \text{for } \emptyset \neq S \subset N. \end{aligned}$$

The restriction of the multiinformation function M to $\mathcal{U}$ will be denoted by m.

Remark Multiinformation generalizes the well-known information-theoretical concept of mutual information and thus it serves as a quantitative characteristic of level of stochastic dependence of more than two random variables. This view led us to accept the name "multiinformation" in [22]. Another name "entaxy" was used in [9].

Now, some properties of these functions are mentioned. They explain why these functions are good tools for study of CI. Firstly, a simple computation gives:

$$M(S) = -H(S) + \sum_{i \in S} H(i) \qquad \text{whenever } S \subset N \qquad (\text{of course } \sum_{i \in \emptyset} H(i) = 0) \tag{3.1}$$

Moreover, it is shown in [22] §4,5:

$$M(ABC) + M(C) \geq M(AC) + M(BC) \qquad \text{whenever } \langle A,B,C \rangle \in T(N) \tag{3.2}$$

and

$$M(ABC) + M(C) = M(AC) + M(BC) \;\text{ iff }\; A \perp B | C \;(P) \quad \text{for } \langle A,B,C \rangle \in T(N) \tag{3.3}$$

The idea of application of multiinformation function was the main step in the proof of validity of new properties of CI in [22] and [23]. Nevertheless, the connection of entropic and multiinformation functions with CI was recognized earlier - see [9] and [13].

As concerns the ability to describe CI for probability measures with finite domain the entropic and multiinformation functions are equivalent (see below). Nevertheless, their further capacities differ. Entropic function can be also used to describe functional dependencies hidden in a discrete probability measure [11] while multiinformation function can be applied to study of CI for continous or "mixed" probability measures [22].

Lemma 1
Let P be a probability measure over N and u be an imset. Then the following equalities are equivalent:
(a) $\langle m, u\rangle = 0$
(b) $\langle M, \overline{u}\rangle = 0$
(c) $\langle H, \overline{u}\rangle = 0$.

Proof: (a)$\Leftrightarrow$(b) is evident as $M(S) = 0$ for $S \in exp\, N\backslash\mathcal{U}$. To see (b)$\Leftrightarrow$(c) simply write using (3.1) and (N.2): $\langle M, \overline{u}\rangle = \sum_{S\subset N} M(S)\cdot\overline{u}(S) = \sum_{S\subset N}\{-H(S) + \sum_{i\in S} H(i)\}\cdot\overline{u}(S) =$
$= -\sum_{S\subset N} H(S)\cdot\overline{u}(S) + \sum_{S\subset N}\sum_{i\in S} H(i)\cdot\overline{u}(S) = -\langle H, \overline{u}\rangle + \sum_{j\in N}\sum_{T\subset N, j\in T} H(j)\cdot\overline{u}(T) =$
$= -\langle H, \overline{u}\rangle + \sum_{j\in N} H(j)\cdot\{\sum_{T, j\in T}\overline{u}(T)\} = -\langle H, u\rangle.$ ■

Def. 8 (probability measure complies with imset)
Let P be a probability measure over N and u an imset on $\mathcal{U}$. Say that P *complies with* u iff any of the conditions (a) - (c) from Lemma 1 is fulfilled.

The above defined concept is related to the product formula validity as follows:

Lemma 2
Let P be a probability measure over N and u be an imset on $\mathcal{U}$. Consider the following conditions:
(a) P satisfies the product formula given by u
(b) $\forall\, x \in \prod_{i\in N} X_i$ with $P(x) > 0$ it holds: $\prod_{S\subset N} P^S(x_S)^{\overline{u}(S)} = 1$
(c) P complies with u.
Then (a) $\Rightarrow$ (b) $\Rightarrow$ (c).

Proof: (a)$\Rightarrow$(b) is evident. To show (b)$\Rightarrow$(c) consider some $x \in \prod_{i\in N} X_i$ with $P(x) > 0$ and write using properties of logarithm:
$\sum_{S\subset N}\overline{u}(S)\cdot ln\, P^S(x_S) = \sum_{S\subset N} ln\, P^S(x_S)^{\overline{u}(S)} = ln\prod_{S\subset N} P^S(x_S)^{\overline{u}(S)} = ln\, 1 = 0.$
By multiplying these equalities by $P(x)$ and summing over all such $x \in \prod_{i\in N} X_i$ get :
$0 = \sum_{x, P(x)>0} P(x)\cdot\sum_{S\subset N}\overline{u}(S)\cdot ln\, P^S(x_S) = \sum_{x, P(x)>0}\sum_{S\subset N} P(x)\cdot\overline{u}(S)\cdot ln\, P^S(x_S) =$
$= \sum_{S\subset N}\overline{u}(S)\cdot\sum_{x, P(x)>0} P(x)\cdot ln\, P^S(x_S) = \sum_{S\subset N} -\overline{u}(S)\cdot H(S) = -\langle H, \overline{u}\rangle$
, i.e. the condition (c) from Lemma 1 holds. ■

4. STRUCTURAL IMSETS

The class of structural imsets is introduced in this section and the corresponding dependency model is defined for every such imset. Then it is shown that complying a probability measure with a structural imset introduced in the preceding section can be interpreted as partial description of the CI-structure (namely by means of the corresponding dependency model). Special attention is devoted to the question how to recognize structural imsets. Last result says that any possible model of CI-structure can be completely described in such a way by a structural imset.

The class of all imsets on $\mathcal{U}$ is too wide for our purposes. Certain subclass will be used to describe CI-structures. These imsets, called structural, can be introduced as "combinations" of so-called elementary imsets defined below.

Def. 9 (elementary imset)
An imset $u \in Z(\mathcal{U})$ is called *elementary* iff its natural extension has the form:
$\overline{u} = \delta_{S\cup T} - \delta_S - \delta_T + \delta_{S\cap T}$ where $S, T \subset N$ $card\, S\setminus T =$ $card\, T\setminus S = 1$.
The set of elementary imsets will be denoted by E.

The following example illustrates this concept in case $card\, N = 4$.

Example 2
Suppose that $N = \{1, 2, 3, 4\}$. By definition every elementary imset is "produced" by a couple $[S, T]$, necessarily $1 \le card\, S = card\, T \le card\, N - 1$.
Thus, elementary imsets can be naturally divided into classes according to the cardinality of "producing" sets. In the considered case three classes can be distinguished:

$E_1 \ldots$ i.e. $card\, S = card\, T = 1$
for instance $S = \{1\}$ and $T = \{2\}$ gives $\overline{u} = \delta_{\{1,2\}} - \delta_{\{1\}} - \delta_{\{2\}} + \delta_\emptyset$
and hence $u = \delta_{\{1,2\}}$.
The corresponding list follows :
$\delta_{\{1,2\}}, \delta_{\{1,3\}}, \delta_{\{2,3\}}, \delta_{\{1,4\}}, \delta_{\{2,4\}}, \delta_{\{3,4\}}$.

$E_2 \ldots$ i.e. $card\, S = card\, T = 2$
for instance $S = \{1, 2\}$ and $T = \{2, 3\}$ gives $\overline{u} = \delta_{\{1,2,3\}} - \delta_{\{1,2\}} - \delta_{\{2,3\}} + \delta_{\{2\}}$
and hence $u = \delta_{\{1,2,3\}} - \delta_{\{1,2\}} - \delta_{\{2,3\}}$.
The corresponding list follows :
$\delta_{\{1,2,3\}} - \delta_{\{1,2\}} - \delta_{\{2,3\}}, \delta_{\{1,2,3\}} - \delta_{\{1,2\}} - \delta_{\{1,3\}}, \delta_{\{1,2,3\}} - \delta_{\{1,3\}} - \delta_{\{2,3\}}$,
$\delta_{\{1,2,4\}} - \delta_{\{1,2\}} - \delta_{\{2,4\}}, \delta_{\{1,2,4\}} - \delta_{\{1,2\}} - \delta_{\{1,4\}}, \delta_{\{1,2,4\}} - \delta_{\{1,4\}} - \delta_{\{2,4\}}$,
$\delta_{\{1,3,4\}} - \delta_{\{1,3\}} - \delta_{\{1,4\}}, \delta_{\{1,3,4\}} - \delta_{\{1,3\}} - \delta_{\{3,4\}}, \delta_{\{1,3,4\}} - \delta_{\{1,4\}} - \delta_{\{3,4\}}$,
$\delta_{\{2,3,4\}} - \delta_{\{2,3\}} - \delta_{\{2,4\}}, \delta_{\{2,3,4\}} - \delta_{\{2,3\}} - \delta_{\{3,4\}}, \delta_{\{2,3,4\}} - \delta_{\{2,4\}} - \delta_{\{3,4\}}$.

$E_3 \ldots$ i.e. $card\, S = card\, T = 3$
for instance $S = \{1, 2, 3\}$ and $T = \{1, 2, 4\}$ gives
$u = \overline{u} = \delta_N - \delta_{\{1,2,3\}} - \delta_{\{1,2,4\}} + \delta_{\{1,2\}}$.
The corresponding list follows :
$\delta_N - \delta_{\{1,2,3\}} - \delta_{\{1,2,4\}} + \delta_{\{1,2\}}$,
$\delta_N - \delta_{\{1,2,3\}} - \delta_{\{1,3,4\}} + \delta_{\{1,3\}}$,
$\delta_N - \delta_{\{1,2,3\}} - \delta_{\{2,3,4\}} + \delta_{\{2,3\}}$,
$\delta_N - \delta_{\{1,2,4\}} - \delta_{\{1,3,4\}} + \delta_{\{1,4\}}$,
$\delta_N - \delta_{\{1,2,4\}} - \delta_{\{2,3,4\}} + \delta_{\{2,4\}}$,
$\delta_N - \delta_{\{1,3,4\}} - \delta_{\{2,3,4\}} + \delta_{\{3,4\}}$.

Thus the total number of elementary imsets is 24 in this case.

It makes no problem to give the formula for total number of elementary imsets: $card\, N \cdot (card\, N - 1) \cdot 2^{card\, N - 3}$.

Hint: the couple S, T can be characterized by the set $(S \setminus T) \cup (T \setminus S)$ of cardinality 2 and by the intersection $S \cap T$ i.e. a subset of the complement.

Def. 10 (structural imset)
An imset $u \in Z(\mathcal{U})$ will be called *structural* iff it holds:

$$\exists n \in \mathbb{N} \quad k_v \in \mathbb{Z}^+ \text{ (for } v \in E) \qquad n \cdot u = \sum_{v \in E} k_v \cdot v \tag{4.1}$$

The following lemma enables to identify CI-statements with structural imsets and to ensure the correctness of further definition.

Lemma Whenever $\langle A, B, C \rangle \in T(N)$ then the imset $u \in Z(\mathcal{U})$ determined by its natural extension $\overline{u} = \delta_{ABC} - \delta_{AC} - \delta_{BC} + \delta_C$ is a structural imset.

Hint: This can be shown by induction according to $card\, AB$. Whenever $card\, AB = 2$ then u is an elementary imset. In case $card\, A \geq 2$ chose $x \in A$ and "extend" $\overline{u}$ by $\pm(\delta_{ABC \setminus \{x\}} - \delta_{AC \setminus \{x\}})$, however in case $card\, A = 1$ take $x \in B$.

Def. 11 (dependency model corresponding to imset)
a) To every triplet $\langle A, B, C \rangle \in T(N)$ assign the structural imset denoted by $i(\langle A, B, C \rangle)$ and specified by its natural extension $\delta_{ABC} - \delta_{AC} - \delta_{BC} + \delta_C$.
b) Let u be a structural imset. The *dependency model corresponding to* u denoted by I_u is defined as follows:
$\langle A, B, C \rangle \in I_u$ iff [$\exists n \in \mathbb{N} \;\; n \cdot u - i(\langle A, B, C \rangle)$ is a structural imset].

Remark that dependency models corresponding to structural imsets are called *structural semigraphoids* in [25].

Lemma 3
Let P be a probability measure over N and u a structural imset on $\mathcal{U}$. Then the following two conditions are equivalent:
(a) P complies with u
(b) I_u is a submodel of CI-structure of P.

Proof: Recall that M is the multiinformation function for P and m its restriction to $\mathcal{U}$.
I. $\langle m, v \rangle \geq 0$ whenever v is a structural imset.
By (3.2) the inequality holds for elementary imsets, then use (4.1).
II. $\langle m, u \rangle = 0 \Rightarrow$ (b).
Consider $\langle A, B, C \rangle \in I_u$, take the structural imset $n \cdot u - i(\langle A, B, C \rangle)$ with $n \in \mathbb{N}$ and write: $0 = \langle m, n \cdot u \rangle = \langle m, n \cdot u - i(\langle A, B, C \rangle) \rangle + \langle m, i(\langle A, B, C \rangle) \rangle$.
Owing to I. both terms on the right-hand side are nonnegative and therefore they vanish. Thus $0 = \langle m, i(\langle A, B, C \rangle) \rangle = \langle M, \overline{i(\langle A, B, C \rangle)} \rangle$ gives by (3.3) $A \perp B | C \; (P)$.

III. (b) $\Rightarrow$ $\langle m, u\rangle = 0$.
By Def. 10 write $n \cdot u = \sum_{v \in E} k_v \cdot v$ with $n \in \mathbb{N}$, $k_v \in \mathbb{Z}^+$. Clearly, it suffices to show $\langle m, v\rangle = 0$ for each $v \in E$ with $k_v > 0$. For this purpose find $\langle A, B, C\rangle \in T(N)$ with $v = i(\langle A, B, C\rangle)$. By Def. 11 $\langle A, B, C\rangle \in I_u$ and by (b) get $A \perp B|C$ (P); hence by (3.3) derive $\langle m, v\rangle = \langle M, \overline{v}\rangle = 0$. ■

The following question arises in connection with computer implementation of structural imsets: how to recognize whether an imset is structural? The presented definition of structural imset is not suitable for solving this problem. Nevertheless, structural imsets can be characterized in another more appropriate way. To formulate it some concept has to be introduced.

<u>Def. 12</u> (completely convex set function)
A set function $c : \mathcal{U} \to \mathbb{R}$ is called a *completely convex set function* iff
its settled extension $\underline{c}$ (i.e. $\underline{c}(T) = 0$ for $T \in exp\, N \setminus \mathcal{U}$) satisfies the convexity condition:
$\underline{c}(K \cup L) + \underline{c}(K \cap L) \geq \underline{c}(K) + \underline{c}(L)$ whenever $K, L \subset N$.

Remark The adjective 'convex' is borrowed from game theory [14] while the adverb 'completely' indicates that the convexity condition concerns the extension.

<u>Assertion 1</u>
a) Let C denotes the class of completely convex set functions. Whenever u is an imset, then it holds: [u is structural] iff [$\forall c \in C$ $\langle c, u\rangle \geq 0$].
Moreover, C is the largest class satisfying the previous condition.
b) There exists the least finite set of <u>normalized</u> imsets A such that for each imset $u \in Z(\mathcal{U})$ it holds: [u is structural] iff [$\forall a \in A$ $\langle a, u\rangle \geq 0$].
(According to the first part necessarily $A \subset C$.)

Proof: The first part of previous assertion is proved in [25] as Theorem 2.4b. The second part is also mentioned in [25] as Assertion 1.4a, but the essential proof is in [24], Proposition 7b. ■

Thus from the theoretical point of view a clear criterion to recognize a structural imset u is given: simply to check the validity of all inequalities $\langle a, u\rangle \geq 0$ for $a \in A$.
The following result says that structural imsets can describe all possible CI-structures:

<u>Assertion 2</u>
Whenever P is a probability measure over N and I the model of CI-structure of P then there exists a structural imset u such that I corresponds to u.

Proof: see Consequence 2.9 in [25]. ■

Remarks
a) It may happen that different structural imsets have the same corresponding dependency model. Nevertheless, the pertinent equivalence of structural imsets can be grasped

by means of the set A from Assertion 1b, for details see [25].
b) Our original conjecture that the models of CI-structures coincide with the dependency models corresponding to structural imsets appeared unfortunately untrue (see [25]).
c) However, the theory developed in [25] seems to admit modifications which promise to give fitting description of CI-structures for some special "nice" subclasses of probability measures.

5. EQUIVALENCE RESULT

The aim of this paper is to show that a probability measure complies with a structural imset u iff it satisfies the product formula given by u. This result can be shown under certain formal additional assumption on u, called regularity. All structural imsets in case $card\,N \leq 4$ are shown to be regular; we conjecture that every structural imset satisfies this condition. The main theorem contains the desired equivalence result.

Def. 13 (regular structural imset)
Consider a structural imset u and put:
$\mathcal{A}_u = \{S \subset N;\ S \subset T$ for some $T \subset N$ with $\overline{u}(T) < 0\}$
$\mathcal{B}_u = \{S \subset N;\ S \subset T$ for some $T \subset N$ with $\overline{u}(T) > 0\}$.
Say that u is *regular* iff only $\mathcal{E} \subset \mathcal{B}_u$ satisfying the following three conditions:
$[a]$ $\mathcal{E}$ is heriditary (i.e. $K \subset L \in \mathcal{E} \;\Rightarrow\; K \in \mathcal{E}$)
$[b]$ $\mathcal{A}_u \subset \mathcal{E}$
$[c]$ whenever $K, L \in \mathcal{E}$ with $\langle K\backslash L, L\backslash K, K \cap L\rangle \in I_u$ then $K \cup L \in \mathcal{E}$
is $\mathcal{B}_u$ itself.

Example 3
Every $u \in Z(\mathcal{U})$ of the form $i(\langle A, B, C\rangle)$ for $\langle A, B, C\rangle \in T(N)$ is a regular structural imset. Especially, every elementary imset is regular.
Indeed: u is a structural imset according to the lemma before Def. 11. Clearly $\mathcal{A}_u = \{K \subset N; K \subset AC$ or $K \subset BC\}$ and $\mathcal{B}_u = \{K \subset N; K \subset ABC\}$. Supposing $\mathcal{E} \subset \mathcal{B}_u$ satisfies $[a]-[c]$, by $[b]$ get $AC, BC \in \mathcal{E}$. As $\langle A, B, C\rangle \in I_u$ (see Def. 11) by $[c]$ derive $ABC \in \mathcal{E}$ and hence by $[a]$ $\mathcal{B}_u \subset \mathcal{E}$.

Some facts concerning the classes $\mathcal{A}_u$ and $\mathcal{B}_u$ (for a structural imset u) are needed to derive certain sufficient condition for regularity. Firstly, considering $S \in \mathcal{A}_u$ find maximal $T \subset N$ with $[S \subset T\ \&\ \overline{u}(T) \neq 0]$. Denoting $r^T = \sum\{\delta_K; T \subset K\}$ it makes no problem to see $\overline{u}(T) = \langle r^T, \overline{u}\rangle \geq 0$. (for example Assertion 1a resp. $(N.1)-(N.2)$). Hence:

$$\mathcal{A}_u \subset \mathcal{B}_u \tag{5.1}$$

Moreover, evident facts

$$\mathcal{A}_{l\cdot u} = \mathcal{A}_u \quad \mathcal{B}_{l\cdot u} = \mathcal{B}_u \quad I_{l\cdot u} = I_u \text{ whenever } l \in \mathbb{N} \tag{5.2}$$

imply that u is regular iff $l \cdot u$ is regular. To show

$$[u = y + w \quad y, w \text{ structural imsets}] \Rightarrow \mathcal{B}_u = \mathcal{B}_y \cup \mathcal{B}_w \tag{5.3}$$

take $S \in \mathcal{B}_y$ and find a maximal $T \subset N$ with $[S \subset T\ \&\ \overline{y}(T) \neq 0]$. As $\overline{y}(T) > 0$ the hypothesis $S \notin \mathcal{B}_u$ leads to the contradiction: $0 < \langle r^T, \overline{y}\rangle + \langle r^T, \overline{w}\rangle = \langle r^T, \overline{u}\rangle \leq 0$.

Thus $\mathcal{B}_y \subset \mathcal{B}_u$, similarly $\mathcal{B}_w \subset \mathcal{B}_u$ and the inclusion $\mathcal{B}_u \subset \mathcal{B}_y \cup \mathcal{B}_w$ is evident.
You can derive from (5.3) and (5.2) by putting $y = i(\langle K\backslash L, L\backslash K, K \cap L\rangle)$ (see Def. 11):
$\langle K\backslash L, L\backslash K, K \cap L\rangle \in I_u \Rightarrow K \cup L \in \mathcal{B}_u$ (5.4)

Lemma 4 Suppose that every structural imset u satisfies:
$[n \cdot u = \sum_{v\in G} k_v \cdot v$ with $n \in \mathbb{N}$ $\emptyset \neq G \subset E$ $k_v \in \mathbb{N}] \Rightarrow [\exists w \in G \; \mathcal{A}_w \subset \mathcal{A}_u]$ (5.5)
Then every structural imset is regular.

Proof: Prove the regularity of a structural imset u by induction according to
$r = min\,\{g\,;\, \exists G \subset E \; card\,G = g$ such that $n \cdot u = \sum_{v\in G} k_v \cdot v$ for $n \in \mathbb{N}$, $k_v \in \mathbb{Z}^+\}$.
In case $r \leq 1$ either $u = 0$ (then $\mathcal{B}_u = \emptyset$) or $u = l \cdot v$ for $l \in \mathbb{N}$ and $v \in E$ (then use the procedure from Example 3 combined with (5.2)). In case $r > 1$ consider a concrete "minimal decomposition" $n \cdot u = \sum_{v\in G} k_v \cdot v$ with $G \subset E$, $card\,G = r$, $n \in \mathbb{N}$, $k_v \in \mathbb{Z}^+$. Necessarily $k_v > 0$! Using (5.5) find $w \in G$ with $\mathcal{A}_w \subset \mathcal{A}_u$ and put $y = \sum_{v\in G\backslash\{w\}} k_v \cdot v$ i.e. $n \cdot u = y + k_w \cdot w$. To verify the regularity of u consider $\mathcal{E} \subset \mathcal{B}_u$ satisfying $[a] - [c]$ (see Def. 13). Our aim is to show $\mathcal{E} = \mathcal{B}_u$. Owing to (5.3) it suffices to verify $\mathcal{B}_w \subset \mathcal{E}$ and $\mathcal{B}_y \subset \mathcal{E}$.
I. $\mathcal{B}_w \subset \mathcal{E}$
Indeed: Let be $\overline{w} = \delta_{K\cup L} - \delta_K - \delta_L + \delta_{K\cap L}$; as $\mathcal{A}_w \subset \mathcal{A}_u$ by $[b]$ derive $K, L \in \mathcal{E}$ and as $\langle K\backslash L, L\backslash K, K \cap L\rangle \in I_u$ by $[c]$ $K \cup L \in \mathcal{E}$. Hence by $[a]$ $\mathcal{B}_w \subset \mathcal{E}$.
II. $\mathcal{B}_y \subset \mathcal{E}$
Indeed: The desired condition is equivalent to $\mathcal{E} \cap \mathcal{B}_y = \mathcal{B}_y$. Since y is regular by the induction assumption it suffices to verify for $\mathcal{E}' = \mathcal{E} \cap \mathcal{B}_y$ the following three conditions:
$[a']$ $\mathcal{E}'$ is heriditary
$[b']$ $\mathcal{A}_y \subset \mathcal{E}'$
$[c']$ $[K, L \in \mathcal{E}' \quad \langle K\backslash L, L\backslash K, K \cap L\rangle \in I_y] \;\Rightarrow\; K \cup L \in \mathcal{E}'$.
As $\mathcal{B}_y$ is heriditary, $[a']$ follows from $[a]$. It makes no problem to see $\mathcal{A}_y \subset \mathcal{A}_u \cup \mathcal{B}_w$ and hence by $[b]$ and I. $\mathcal{A}_y \subset \mathcal{E}$. Thus $[b']$ follows from (5.1). To show $[c']$ realize that $I_y \subset I_u$ and $[c]$ can be used to derive $K \cup L \in \mathcal{E}$. By (5.4) get $K \cup L \in \mathcal{E}'$. ■

Consequence 1
In case $card\,N \leq 4$ every structural imset is regular.

Proof: The condition (5.5) will be verified for every structural imset in case $card\,N = 4$ (the same method can be used in simpler cases $card\,N = 3$ and $card\,N = 2$). Consider a concrete "decomposition":
$n \cdot u = \sum_{v\in E} k_v \cdot v$ with $n \in \mathbb{N}$, $k_v \in \mathbb{Z}^+$, $G = \{v \in E, k_v > 0\}$ (define $k_v = 0$ for $v \in E\backslash G$).
Divide E into three classes E_1, E_2, E_3 (see Example 2) and put: $p_i = \sum_{v\in E_i} k_v$ for $i = 1, 2, 3$.
Three basic cases can be distinguished:
I. $p_1 = p_2 = 0$
In this case $n \cdot u = \sum_{v\in E_3} k_v \cdot v$. It suffices to take arbitrary $w \in E_3$ with $k_w > 0$ (note

that $w(S) < 0$ & $card\, S = 3$ implies $n \cdot u(S) < 0$ as $v(S) \leq 0$ for every $v \in E_3$).
II. $p_1 = 0 \;\; p_2 > 0$
In this case $n \cdot u = \sum_{v \in E_2 \cup E_3} k_v \cdot v$ and there exist $v \in E_2$ with $k_v > 0$. Consider the case $\mathcal{A}_v \backslash \mathcal{A}_u \neq \emptyset$ (otherwise put $w = v$) and find $S \subset N$ with $v(S) < 0$ and $[\forall T \supset S \;\; u(T) \geq 0]$. More concretely let $S = \{a, b\}$ $v = \delta_{\{a,b,c\}} - \delta_{\{a,b\}} - \delta_{\{a,c\}}$ where $N = \{a, b, c, d\}$. Moreover consider the following elementary imsets:

$$\begin{aligned} w = & \;\; \delta_{\{a,b,d\}} - \delta_{\{a,d\}} - \delta_{\{b,d\}} \\ x = & \;\; \delta_N - \delta_{\{a,b,d\}} - \delta_{\{a,c,d\}} + \delta_{\{a,d\}} \\ y = & \;\; \delta_N - \delta_{\{a,b,d\}} - \delta_{\{b,c,d\}} + \delta_{\{b,d\}} \\ z = & \;\; \delta_{\{a,b,c\}} - \delta_{\{a,b\}} - \delta_{\{b,c\}} \end{aligned}$$

Our aim is to show that $k_w > 0$ and $\mathcal{A}_w \subset \mathcal{A}_u$. For this purpose write $n \cdot [u(\{a, b\}) + u(\{a, b, d\})] = +k_w - k_v - k_x - k_y - k_z$ (for all remaining $t \in E_2 \cup E_3 \;\; t(\{a, b\}) + t(\{a, b, d\}) = 0$). Hence $k_w \geq k_v + k_x + k_y$ implies both $k_w \geq k_v > 0$ and $0 > -k_v \geq k_x - k_w \geq n \cdot u(\{a, d\})$ and $0 > -k_v \geq k_y - k_w \geq n \cdot u(\{b, d\})$ (the inequality $n \cdot u(\{a, d\}) \leq k_x - k_w$ follows from the fact $t(\{a, d\}) \leq 0$ for remaining $t \in E_2 \cup E_3$, similarly $n \cdot u(\{b, d\}) \leq k_y - k_w$).
III. $p_1 > 0$
In this case take arbitrary $w \in E_1$ with $k_w > 0$. By (5.1) and (5.3) $\mathcal{A}_w \subset \mathcal{B}_w \subset \mathcal{B}_u$ i.e. for each $S \in \mathcal{A}_w$ ($card\, S = 1$) get $\sum_{S \subset K} \overline{u}_+(K) > 0$. Nevertheless by $(N.2)$ $\sum_{S \subset K} \overline{u}_-(K) = \sum_{S \subset K} \overline{u}_+(K)$ and hence $S \in \mathcal{A}_u$. ■

So far, we have no example of nonregular structural imset. Our conjecture is that it cannot be found:

Conjecture Every structural imset is regular.

Now, the main result can be proved.

THEOREM
Let P be a probability measure over N and u be a regular structural imset on $\mathcal{U}$. Then the following conditions are equivalent:

(a) P satisfies the product formula given by u (see Def. 4)

(b) $\forall x \in \prod_{i \in N} X_i$ with $P(x) > 0$ it holds $\prod_{S \subset N} P^S(x_S)^{\overline{u}(S)} = 1$

(c) P complies with u (see Def. 8)

(d) I_u is a submodel of CI-structure of P (see Def. 6,11).

Proof: By Lemma 2 (a)⇒(b)⇒(c), by Lemma 3 (c)⇒(d). It remains to show (d)⇒(a). For fixed $x \in \prod_{i \in N} X_i$ two possibilities can occur:

I. $\prod_{S\subset N} P^S(x_S)^{\overline{u}_-(S)} = 0.$
In this case find $K \subset N$ with $[\overline{u}(K) < 0 \;\; \& \;\; P^K(x_K) = 0]$. By (5.1) there exists T with $[\,K \subset T \;\; \& \;\; \overline{u}(T) > 0\,]$. Necessarily $P^T(x_T) = 0$ and hence $\prod_{S\subset N} P^S(x_S)^{\overline{u}_+(S)} = 0$.
II. $\prod_{S\subset N} P^S(x_S)^{\overline{u}_-(S)} > 0.$
Put $\mathcal{E} = \{S \in \mathcal{B}_u; P^S(x_S) > 0\}$. Evidently $\mathcal{E}$ is heriditary and by the assumption $\mathcal{A}_u \subset \mathcal{E}$. Also the condition [c] from Def. 13 is valid, owing to (d):
$\langle K\backslash L, L\backslash K, K\cap L\rangle \in I_u \quad \Rightarrow \quad P^{K\cup L}(x_{K\cup L}) \cdot P^{K\cap L}(x_{K\cap L}) = P^K(x_K) \cdot P^L(x_L)$
and hence $P^{K\cup L}(x_{K\cup L}) > 0$. Therefore the regularity of u implies $\mathcal{B}_u = \mathcal{E}$.
By Def. 10 $\;\; n \cdot u = \sum_{v\in E} k_v \cdot v$ for $n \in \mathbb{N}$, $k_v \in \mathbb{Z}^+$. For every $v \in E$ with $k_v > 0$ consider $\langle A, B, C\rangle \in T(N)$ with $v = i(\langle A, B, C\rangle)$. As $\langle A, B, C\rangle \in I_u$, by (d) $A \perp B|C\,(P)$ and therefore by Def. 5 derive $\prod_{S\subset N} P^S(x_S)^{\overline{v}_+(S)} = \prod_{S\subset N} P^S(x_S)^{\overline{v}_-(S)}$.
All these formulas can be multiplied and therefore it holds:

$$\prod_{S\subset N} P^S(x_S)^{\sum_{v\in E} k_v\cdot\overline{v}_+(S)} = \prod_{S\subset N} P^S(x_S)^{\sum_{v\in E} k_v\cdot\overline{v}_-(S)} \tag{5.6}$$

Put $w = \sum_{v\in E} k_v \cdot \overline{v}_+ - n \cdot \overline{u}_+$, evidently $w = \sum_{v\in E} k_v \cdot \overline{v}_- - n \cdot \overline{u}_-$. Of course $w \geq 0$ and $w(S) > 0$ implies $[\exists v \in E \;\; k_v > 0 \;\; \overline{v}(S) > 0]$ i.e. $S \in \mathcal{B}_v \subset \mathcal{B}_u$ by (5.3) and (5.2). Therefore $S \in \mathcal{E}$ and $P^S(x_S)^{w(S)} > 0$. Together $\prod_{S\subset N} P^S(x_S)^{w(S)} > 0$ and the equality (5.6) can be divided by this number to get

$$\prod_{S\subset N} P^S(x_S)^{n\cdot\overline{u}_+(S)} = \prod_{S\subset N} P^S(x_S)^{n\cdot\overline{u}_-(S)} \tag{5.7}$$

Hence, the desired product formula easily follows. ■

Remark The previous proof can be easily modified to show that for every strictly positive probability measure and every structural imset u the conditions (a), (b), (c), (d) are equivalent.

CONCLUSIONS

Thus, the theorem above relates three approaches to description of CI-structures:
- by means of dependency models
- by means of imsets (information - theoretical definition)
- by means of product formula validity
and shows their equivalence.

Note that the description by means of imsets (and faces which are behind) is systematically treated and illustrated by examples in [25]. It is endowed by a deductive mechanism allowing to infer CI-statements from an input piece of information about CI-structure (finitely-implementable from theoretical point of view).

The description by means of product formula can be understood as a step to interpretation of these CI-structures. It seems to me that the presented models of description of CI-structures have similar reasons (or rights) to be called explicable as hierarchical log-linear models. In fact, a general log-linear model is specified by certain "formula" for the probability measure, namely expressing it as a product of marginal factors. This is close to the presented formulas and in some special cases (decomposable models) even equivalent.

Another interesting analogy of our product formulas can be found in [10] where so-called functional expressions satisfying the unity sum property are delt with. Some of them (for example the expression from Example 4 there) correspond to product formulas representing CI-structure.

Note that as reported in [25] some graphical descriptions of CI-structures can be "translated" to imsets and these can be "forwarded" to product formulas. To inform the reader we give the corresponding imset expressions here (without proof).

Having an influence diagram (= directed acyclic graph) let $\pi(k)$ denotes the set of parents of a node $k \in N$. The corresponding imset can be given by its natural extension: $\overline{u} = \delta_N - \delta_\emptyset + \sum_{k \in N} \left\{ \delta_{\pi(k)} - \delta_{\{k\} \cup \pi(k)} \right\}$.

Having a decomposable model specified by a triangulated (undirected) graph let $\mathcal{C} \subset exp\ N$ denote the class of its maximal cliques. The corresponding imset is specified by: $\overline{u} = \delta_{\cup \mathcal{C}} + \sum_{\emptyset \neq \mathcal{B} \subset \mathcal{C}} (-1)^{card\, \mathcal{B}} \cdot \delta_{\cap \mathcal{B}}$
($\cup \mathcal{B}$ resp. $\cap \mathcal{B}$ denotes the union resp. intersection of sets from $\mathcal{B}$).

Acknowledgements

I am indebted to Francesco Malvestuto; one of his comments during his visit in Prague last year helped me to overcome certain inconvenience. I am also grateful to my colleague František Matúš for stimulating discussion.

REFERENCES

1. M.Aigner, Combinatiorial Theory, Springer Verlag, Berlin-Heidelberg-New York, 1979.
2. J.N.Darroch and S.L.Lauritzen and T.P.Speed, Markov fields and log-linear interaction models for contingency tables, Annals of Statistics 8 (1980) 522-539.
3. A.P.Dawid, Conditional independence in statistical theory (with discussion), Journal of the Royal Statistical Society ser. B 41 (1979) 1-31.
4. D.Geiger and J.Pearl, Logical and algorithmic properties of conditional independence and graphical models, to appear in Annals of Statistics (1993?).
5. R.A.Howard and J.E.Matheson, Influence diagrams, in Principles and Applications of Decision Analysis 2, Strategic Decisions Group, Menlo Park, California, 1984.
6. V.Isham, An introduction to spatial point processes and Markov random fields, International Statistical Review 49 (1981) 21-43.
7. H.Kiiveri and T.P.Speed and J.B.Carlin, Recursive causal models, Journal of Australian Mathematical Society ser.A 36 (1984) 30-52.
8. S.L.Lauritzen and T.P.Speed and K.Vijayan, Decomposable graphs and hypergraphs, Journal of Australian Mathematical Society 36 (1984) 12-29.
9. F.M.Malvestuto, Theory of random observables in relational data bases, Information Systems 8 (1983) 281-289.
10. F.M.Malvestuto, Existence of extensions and product extensions for discrete probability distributions, Discrete Mathematics 69 (1988) 61-77.

11. F.Matúš, Probabilistic conditional independence structures and matroid theory: backgrounds, in Proceedings of WUPES 1991 (Workshop on Uncertainty Processing in Expert Systems), September 9-12, 1991, Alšovice, Czechoslovakia.

12. J.Moussouris, Gibbs and Markov random systems with constraints, Journal of Statistical Physics 10 (1974) 11-33.

13. K.K.Nambiar, Some analytic tools for design of relational database systems, in Proceedings of Conference on Very Large Data Bases 1980.

14. J.Rosenmüller and H.G.Weidner, Extreme convex set functions with finite carrier: general theory, Discrete Mathematics 10 (1974) 343-382.

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

16. J.Pearl, Markov and Bayes networks: a comparison of two graphical representations of probabilistic knowledge, technical report CSD-860024, R-46-1, UCLA, Los Angeles, 1986.

17. J.Pearl and D.Geiger and T.Verma, The logic of influence diagrams, chapter 4 in Influence Diagrams, Belief Nets and Decision Analysis, (R.M. Oliver and J.Q. Smith eds.), John Wiley 1990.

18. R.D.Shachter, Probabilistic inference and influence diagrams, Operations Research 36 (1988) 589-604.

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

20. D.J.Spiegelhalter and S.L.Lauritzen, Sequential updating on conditional probabilities of directed graphical structures, Networks 20 (1990) 579-605.

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

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

23. 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 377-396, Academia, Prague, 1992.

24. M.Studený, Convex cones in finite-dimensional real vector spaces, Kybernetika 29 (1993) (to appear).

25. M.Studený, Description of structures of stochastic conditional independence by means of faces and imsets, series of 3 papers submitted to International Journal of General Systems.

26. S.Ur and A.Paz, The representation power of probabilistic knowledge by undirected graphs and directed acyclic graphs: a comparison, in Proceedings of WUPES 1991 (Workshop on Uncertainty Processing in Expert Systems), September 9-12, 1991, Alšovice, Czechoslovakia.

27. N.Wermuth and S.L.Lauritzen, Graphical and recursive models for contingency tables, Biometrika 70 (1983) 537-552.

28. S.Wright, Correlation and causation, Journal of Agricultural Research 20 (1921) 557-585.