Int. J. General Systems, Vol. 23, pp. 123–137
Reprints available directly from the publisher
Photocopying permitted by license only

DESCRIPTION OF STRUCTURES OF STOCHASTIC CONDITIONAL INDEPENDENCE BY MEANS OF FACES AND IMSETS 1st part: introduction and basic concepts[1]

MILAN STUDENÝ

Institute of Information Theory and Automation, Czech Academy of Sciences, Pod vodárenskou věží 4, 182 08 Prague 8, Czech Republic

(Received September, 1992, in final form 25 April, 1994)

Global Abstract (for all three parts) The work presents a new approach to the mathematical description of stochastic conditional independence structures of a finite number of random variables. The new approach is related to the classical approaches, that is to the use of directed acyclic graphs (Bayesian networks), undirected graphs (Markov networks) and dependency models (semigraphoids). The approach provides a deductive mechanism to infer probabilistically valid consequences of positive information about conditional independence structure. This mechanism is much more powerful than the use of semigraphoids as it includes, from the classical point of view, an infinite number of inference rules. Nevertheless, from the theoretical point of view, it is finitely implementable. The developed theory is illustrated by examples showing how it is applicable.

INDEX TERMS[2]: Conditional independence, dependency model, imset, scalar product ordering, base, skeleton.

INTRODUCTION

History of Description of Conditional Independence Structures

Attempts to describe adequately the structure of dependence (or independence) relationships among random variables have a long and rich tradition. Graphs (both directed and undirected) were used to make these structures understandable by humans (this was probably started by geneticist S. Wright [1934]). Directed acyclic graphs form the basis of the method of influence diagrams [Howard and Matheson, 1981], [Shachter, 1986], [Smith, 1989], which is widely used in artificial intelligence—see [NETWORKS, 1990]. Similarly, but in a slightly different way (as reported in [Smith, 1989] p. 661) directed acyclic graphs are used in recursive models for contingency tables [Wermuth and Lauritzen, 1983], [Kiiveri, *et al.*, 1984]. Use of undirected graphs stems from Markov field theory [Moussouris, 1974], [Darroch, *et al.*, 1980], [Lauritzen, *et al.*, 1984].

[1]This research was supported by the internal grants of Czech Academy of Sciences n. 27510 "Explanatory power of probabilistic expert systems: theoretical background" and n. 27564 "Knowledge derivation for probabilistic expert systems".

[2]AMS classification: 68T 30, 62B 10.

The concept that was gradually recognized as underlying these methods is stochastic *conditional independence* (CI). This concept was accentuated in modern statistics first by Dawid [1979], and later by other researchers [Mouchart and Rolin, 1984], [van Putten and van Shuppen, 1985]. Spohn [1980] studied the concept of CI from a philosophical point of view. And finally, its importance for probabilistic expert systems was explicitly discerned and highlighted by Pearl and Paz [1987], who introduced the concept of a *dependency model* [Pearl, 1988] to describe the structure of multidimensional probability measures (and many other structures too). Roughly speaking, a dependency model is a set of statements of the form $A \perp B \mid C$ where A, B, C are disjoint subsets of the set of attributes (variables, symptoms) N, interpreted in probabilistic setting as $[\xi_i]_{i\in A}$ is conditionally independent of $[\xi_i]_{i\in B}$ given $[\xi_i]_{i\in C}$ (here $[\xi_i]_{i\in N}$ is the system of random variables). In contrast to graphical methods, dependency models allow the description of all possible probabilistic CI-structures. They also give direct access to information about CI-structure, since any CI-statement $A \perp B \mid C$ can be interpreted (in harmony with an equivalent definition of CI) as representing a qualitative relationship among attributes, namely: if the values of the attributes from C are known, the information about A becomes irrelevant with respect to the information about B.

However, owing to properties of CI treated earlier by Dawid [1979] (see also [Spohn, 1980], [Smith, 1989]), this conception would be too wide to describe CI-structures. Therefore, a special class of dependency models satisfying the mentioned properties was introduced and named the class of *semigraphoids* [Pearl, 1988] together with a subclass of *graphoids* intended to describe structures of strictly positive probability measures. Semigraphoids were defined as dependency models closed under 4 concrete inference rules (called axioms by Pearl and Paz), graphoids have one additional inference rule. As such they provide an easy deductive mechanism to infer CI-statements as valid consequences of an input list of CI-statements. The original hypothesis from [Pearl and Paz, 1987] that graphoids coincide with dependency models probabilistically representable by a strictly positive measure was refuted in [Studený, 1989a] where a new independent property of stochastic CI was found. Later, it has appeared that probabilistically representable dependency models cannot be characterized as dependency models closed under a finite number of inference rules [Studený, 1992]. On the other hand, such a characterization can be found for certain important substructures of CI-structure—see [Geiger and Pearl, 1993], [Matúš, 1995].

Motivation

Nevertheless, the mentioned results do not exclude the possibility of an efficient deductive mechanism for CI, they only say that the use of inference rules is unsuitable. This led me to an attempt to develop a way of inferring CI-statements from which the mentioned crucial disadvantage of the inference rule approach is removed. In fact, I looked for a new approach to the description of CI-structures. The first attempt in this direction was made in [Studený, 1989b] but that original proposal was immature. I have since made an effort to clarify this approach and make the theory elegant and computer implementable. The result of this effort is presented here. From a classical point of view the proposed deductive mechanism involves an infinite number of inference rules, in fact, all those rules whose probabilistic validity can be proved by the method used in [Studený, 1992]. The presented theory can be understood either solely as a proposal for a "finitely implementable" deductive mech-

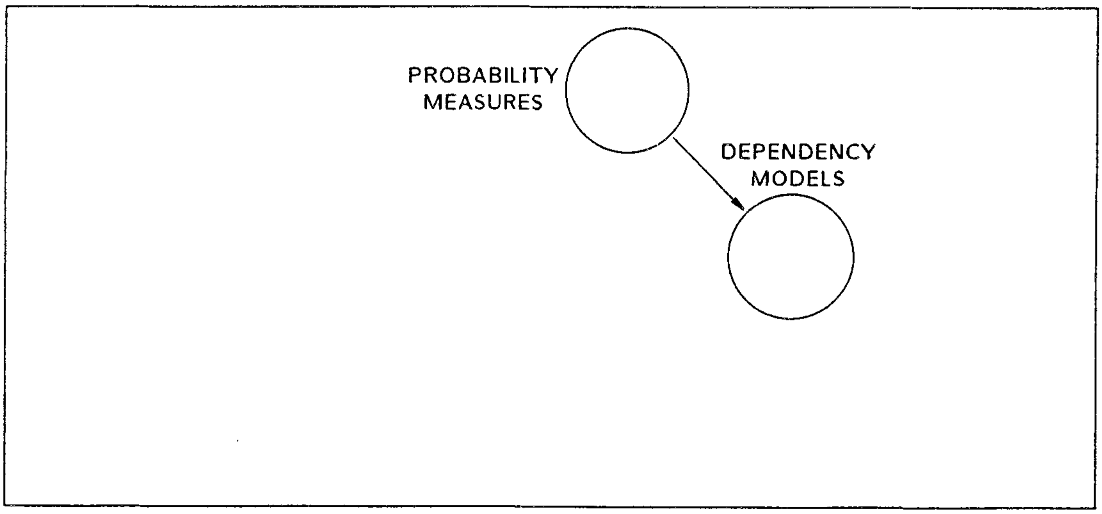

Figure 1.1 A dependency model is assigned to every probability measure.

anism for (probabilistically representable) dependency models or as an alternative way to describe CI-structures.

The fundamental concept of the presented theory is that of a *face*.[3] Formally, a face is defined as a set of so-called *imsets* (i.e. functions ascribing integers to sets of attributes) closed under three operations named *nontriviality*, *composition* and *decomposition*. Here, faces are used to represent CI-structures. They can be understood as a counterpart of dependency models. That is, imsets correspond to CI-statements, faces (i.e. sets of imsets) to dependency models (i.e. sets of CI-statements) and the three above mentioned operations to inference rules. Our problem is to represent faces in a computer. Two ways are proposed. The first one uses so-called *generating imsets* to represent faces. The second one, which uses so-called *portraits*, seems to be more fitting but it requires a complete list of certain imsets, which is named the *skeleton*.

The whole work is organized as follows. It is divided into three parts. The first part is contained in this paper; in addition to this Introduction it defines some basic concepts. The second part is devoted to the mathematical foundations of the theory. The third part gives several examples of its use. Every result or definition is identified by two numbers, the first indicating the part where it can be found and the second its location within that part.

Construction of the Theory

We will now outline the construction of the theory more concretely and by means of illustrative diagrams. The primary situation is recapitulated (this time with precise mathematical definitions) in section 1.1 where a dependency model is assigned to every probability measure (see Figure 1.1).

[3]The name "face" was motivated by an analogy with the theory of convex polytopes: the class of "our" faces is isomorphic with the lattice of faces (concepts from [Broudsted, 1983]) of a certain cone. However, the motivation behind our nomenclature is immaterial from the point of view of this paper.

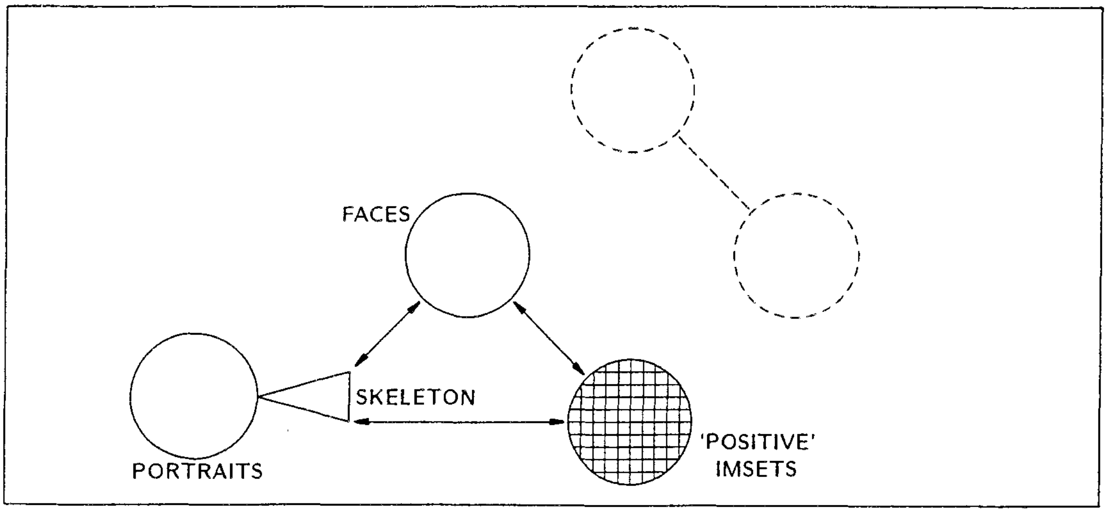

Figure 1.2 Faces are identified with imsets and portraits through the skeleton.

The next two sections contain preparatory mathematics: section 1.2 gives basic definitions concerning imsets and section 1.3 contains some assertions concerning quasiorderings on imsets. The assertions are based on results (of a geometric nature) about cones in finite-dimensional real vector spaces which are proved in another article [Studený, 1993a].

The second installment of the work, subtitled *2nd part: basic theory*, is deeper. In section 2.1 the main theory is described, for the time being, without reference to probability measures or dependency models. In order to make the theory more widely applicable[4] the theory is developed for any so-called scalar product ordering, but reasonable results are achieved only for a special class of orderings called finitely established orderings. Thus, faces with the corresponding deductive mechanism (called *facial implication*) are defined there and every face is identified with certain "positive" imsets (i.e. positive with respect to the considered ordering). This mapping endows the class of "positive" imsets with a corresponding equivalence, called *facial equivalence*. Facial implication can be tested easily by a computer in the case that the *skeleton* is at disposal. Then faces or imsets can be represented by their *portraits* (for an illustrative picture see Figure 1.2).

The "right ordering" (the faces of which are used to represent CI-structures in this work) is introduced in section 2.2 and called *the structural ordering*. Two equivalent definitions of this finitely established ordering are mentioned and the existence of the skeleton is derived.

The next step is to relate structural faces and imsets to dependency models, and this is the subject of the first two subsections of the section 2.3. Namely, a dependency model is ascribed to every structural face and this induces an identification of structural imsets (endowed with the facial equivalence) with dependency models. The ascribed dependency model is called a *structural semigraphoid* (see Figure 1.3).

Finally, the image is completed by assigning a structural face to every probability measure. Of course, the new definition is concordant with the primary approach: the

[4]The reason for this is explained in Conclusions—see the third instalment of the work.

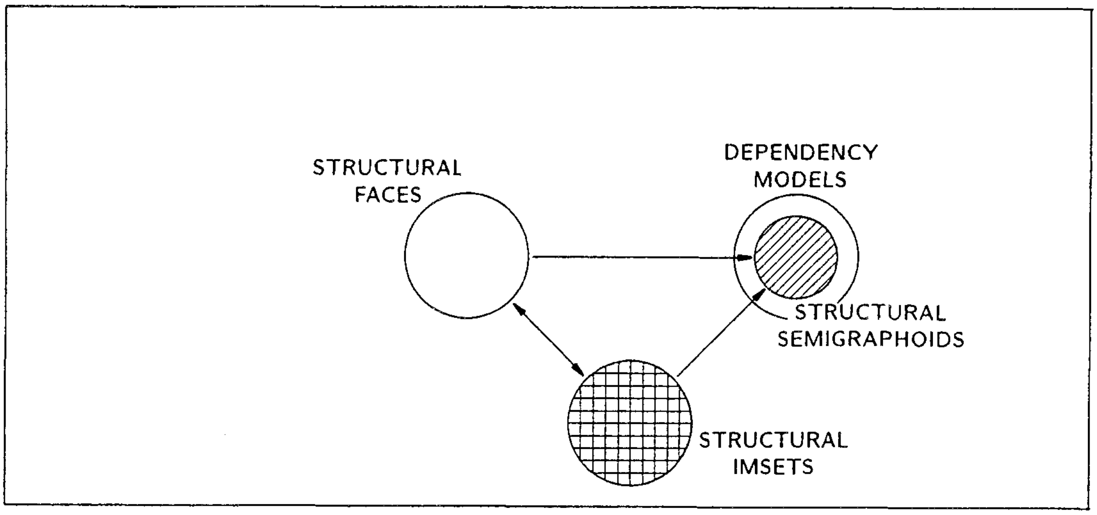

Figure 1.3 A dependency model is assigned to every structural face.

dependency model ascribed to that face coincides with the dependency model assigned originally to the probability measure (see Figure 1.4).

The last installment of the work, subtitled *3rd part: examples of use and appendices*, illustrates the implementation of facial implication and relates the theory to other methods for describing CI-structures. Examples show how to transform information about CI-structure in the form of a dependency model (especially an individual CI-statement), Bayesian network (i.e. directed acyclic graph) and Markov network (i.e. undirected graph) into an imset. Thus, the theory can be applied in those areas as well.

A method of proving the probabilistic soundness of a given inference rule is also illustrated. This procedure is quite simple. The reader who is not interested in mathematics can read and use it almost immediately (only sections 1.1, 1.2 and Definition 2.10 are needed). The proof of a result concerning the structural ordering is shifted

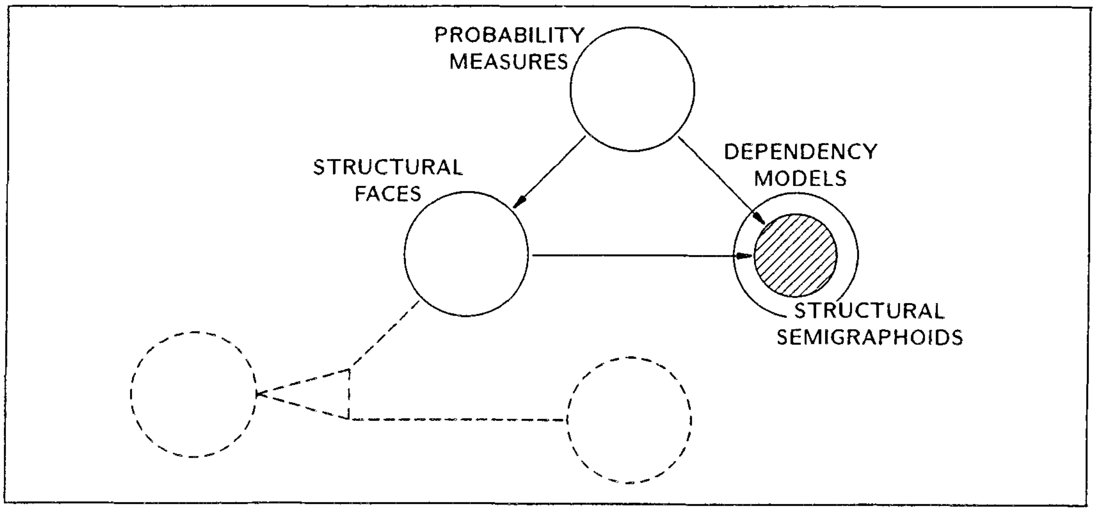

Figure 1.4 A structural face is assigned to every probability measure.

to Appendix A and some remarks concerning the operation of contraction of imsets are contained in Appendix B. The advantages and disadvantages of and prospects for the presented approach are discussed in Conclusions.

NOTATION AND BASIC ARRANGEMENTS

Firstly, we review the symbols for number sets:

$\mathbb{R}$	real numbers
$\mathbb{Q}^+$	nonnegative rational numbers
$\mathbb{Z}$	integers
$\mathbb{Z}^+ = \mathbb{Z} \cap [0, \infty)$	nonnegative integers (including zero)
$\mathbb{N} = \{1, 2, 3, \ldots\}$	natural numbers (= positive integers).

Throughout this paper we will deal with the following situation: A finite set N having at least two elements called the *basic set* is given, i.e. $2 \leq \text{card}\, N < \infty$. The class of all its subsets will be denoted by exp N. The class of *nontrivial* subsets of N, i.e. subsets having at least two elements will be denoted by $\mathcal{U}$:

$$\mathcal{U} = \{S \subset N;\ \text{card}\, S \geq 2\}.$$

Having a set $T \subset N$, its *indicator* i.e. the zero-one function on exp N (possibly restricted to $\mathcal{U}$), is defined as follows:

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

For disjoint sets $A, B \subset N$ the juxtaposition AB will abbreviate $A \cup B$ in many examples. Finally, symbols for several classes of functions on $\mathcal{U}$ are introduced:

$R(\mathcal{U})$	the class of real functions on $\mathcal{U}$
$Q(\mathcal{U})$	the class of rational-valued functions on $\mathcal{U}$
$Z(\mathcal{U})$	the class of integer-valued functions on $\mathcal{U}$
$Z^+(\mathcal{U})$	the nonnegative integer-valued functions on $\mathcal{U}$.

(Notice that the notation is concordant with the symbols for number sets.)

1.1 DEPENDENCY MODELS

In this section the "classical" approach to the description of CI-structures, namely by means of the concept of dependency model, is recalled. Using this approach it is possible to describe CI-structures by means of graphs (both undirected ones and directed acyclic ones) and also to try to formalize them by means of a simple deductive system. First, elementary concepts (probability measure, conditional independence) are recalled. Then the concept of dependency model is defined and related to CI. Finally, special attention is devoted to semigraphoids—a special type of dependency model.

1.1.1 Probability Measures and Conditional Independence

We begin by introducing the relevant class of probability measures. These measures are distributions of systems of finitely-valued random variables indexed by the basic set N, therefore their domains are cartesian products of finite sets indexed by N. (I follow a custom, common in the literature, by alluding to random variables but in fact dealing with probability measures.)

DEFINITION 1.1 (probability measure over N, marginal measure)
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 $\Pi_{i\in N}X_i$.
The class of all these measures will be denoted by $\mathcal{P}(N)$.
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:

$$P^S(A) = P(A \times \Pi_{i\in N\setminus S} X_i) \quad \text{whenever } A \subset \Pi_{i\in S} X_i.$$

Moreover, P^N is defined as P itself.

Depending on the various authors, elements of N may be called variables or symptoms or attributes, while sets X_i are called their domains. Of course, supposing P is the distribution of a random system $[\xi_i]_{i\in N}$, the marginal measure P^S (where $\emptyset \neq S \subseteq N$) is the distribution of its subsystem $[\xi_i]_{i\in S}$.

The aim is to deal with conditional independence relationships among random variables within a random system $[\xi_i]_{i\in N}$. More precisely, we are considering statements of the type "$[\xi_i]_{i\in A}$ is conditionally independent of $[\xi_i]_{i\in B}$ given $[\xi_i]_{i\in C}$", where A, B, C are three disjoint subsets of N. Surely, such statements are defined through the distribution of $[\xi_i]_{i\in N}$; one of possible definitions follows.

DEFINITION 1.2 (conditional independence)
Suppose that P is a probability measure over N (on $\Pi_{i\in N}X_i$ concretely). Having a triplet $\langle A, B, C\rangle$ of pairwise disjoint subsets of N, where A, B are nonempty, we say that the conditional independence statement (CI-statement) $A \perp B \mid C$ is valid for P or that *P obeys the triplet $\langle A, B, C\rangle$* iff

$$\forall\ [x_i]_{i\in N} \in \prod_{i\in N} X_i \quad P^{ABC}([x_i]_{i\in ABC}) \cdot P^C([x_i]_{i\in C}) = P^{AC}([x_i]_{i\in AC}) \cdot P^{BC}([x_i]_{i\in BC}).$$

($P(a)$ is written instead of $P(\{a\})$ and the convention $P^{\emptyset}([x_i]_{i\in\emptyset}) = 1$ is accepted here.)

Note that there are many equivalent definitions of conditional independence, one of which allows CI to be interpreted as a certain (nonnumerical) relationship among variables or symptoms. This makes it possible to use qualitative (structural) information of this type to form the knowledge base of a probabilistic expert system.

1.1.2 The Concept of Dependency Model

Now, what does it mean to say that two random systems (indexed by N) have the same CI-structure? Naturally, that the corresponding CI-statements have the same true value! The following concept of dependency model formalizes this intuitive notion.

DEFINITION 1.3 (dependency model)
Denote by $T_*(N)$ the set of all *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*.

Remark This concept was introduced (in a slightly different form) by the group around Judea Pearl ([Pearl, 1988], [NETWORKS, 1990] pp. 507–534) as a tool to describe various types of relational structure: stochastic conditional independence, separation in graphs or embedded multivalued dependency (from the theory of relational databases). I accepted and modified this concept in [Studený, 1992].

In this article dependency models are used to describe structures of conditional stochastic independence:

DEFINITION 1.4 (submodel and model of CI-structure)
Suppose that I is a dependency model over N and P a probability measure over N. Then I is a *submodel of the* CI-*structure of* P iff P obeys all triplets belonging to I. Moreover, I is the *model of the* CI-*structure of* P iff I is exactly the set of triplets obeyed by P.

Terminological remark This terminology attempts to highlight the presented view on dependency models. Many other phrases are used in the literature to say that I is the set of triplets obeyed by P: 'I is induced by P' in [Ur and Paz, 1994], 'P is perfect for I' in [Geiger and Pearl, 1993], 'I is probabilistically representable by P' in [Matúš, 1992], 'I is conditional-independence relation corresponding to P' in [Studený, 1992].

DEFINITION 1.5 (probabilistically representable models)
Suppose that $\Phi \subseteq \mathcal{P}(N)$ and I is a dependency model over N. Then I is *represented* in Φ iff there exists $P \in \Phi$ such that I is the model of CI-structure of P.
Moreover, I is called *probabilistically representable* iff it is represented in $\mathcal{P}(N)$.

1.1.3 Semigraphoids

Of course, some dependency models are not probabilistically representable. Well-known properties of conditional independence treated by Dawid [1979] or Spohn [1980] imply that every probabilistically representable dependency model satisfies the following conditions:

DEFINITION 1.6 (semigraphoid)
A dependency model I is called a *semigraphoid* iff it satisfies:

(a) $\langle A, B, C\rangle \in I \Leftrightarrow \langle B, A, C\rangle \in I$ whenever $\langle A, B, C\rangle \in T_*(N)$

(b) $\langle A, BC, D\rangle \in I \Leftrightarrow \langle A, B, CD\rangle, \langle A, C, D\rangle \in I$

whenever $\langle A, B, CD\rangle, \langle A, C, D\rangle \in T_*(N)$.

The concept of semigraphoid was introduced by Pearl [1988] in rather different form. He introduced semigraphoids as dependency models closed under 4 inference rules (called by him axioms):

DEFINITION 1.7 (semigraphoid derivability)
Suppose that $I \subseteq T_*(N)$ and $t \in T_*(N)$. Say that t is *derivable from I* (by semigraphoid rules) and write $I \vdash_{sem} t$ iff there exists a derivation sequence $k_1, \ldots, k_n \subseteq$

$T_*(N)$ where $k_n = t$ and for each k_i either $k_i \in I$ or k_i is a direct consequence of some preceding k_js by virtue of some of the following inference rules:

$\langle A, B, C\rangle \rightarrow \langle B, A, C\rangle$ 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

$[\langle A, C, D\rangle, \langle A, B, CD\rangle] \rightarrow \langle A, BC, D\rangle$ contraction

(Both antecedents and consequents of these rules are assumed to be elements of $T_*(N)$.)

Remark This terminology was probably motivated by an idea to describe special dependency models and therefore CI-structures by means of both undirected and directed acyclic graphs—see Introduction for references.

Evidently, a dependency model I is a semigraphoid iff every triplet derivable by the above mentioned deductive system belongs to I. Nevertheless, the aim of Pearl and Paz's approach was to characterize probabilistically representable dependency models as models closed under a finite number of inference rules, i.e. to formalize probabilistic implication:

DEFINITION 1.8 (probabilistic implication)
Suppose that $I \subseteq T_*(N)$ and $t \in T_*(N)$. Say that *I probabilistically implies t* and write $I \models t$ iff every $P \in \mathcal{P}(N)$ that obeys I (i.e. all triples from I) also obeys t.

Terminological remark Various phrases are used in the literature for this relation, namely 't is logically implied by I' in [Geiger and Pearl, 1990], 't is valid consequence of I' in [Pearl, *et al.*, 1990] or 't is entailed by I' in [Geiger and Pearl, 1993].

1.2 IMSETS

This section introduces some elementary concepts. First, the concepts of imset and multiset are defined together with corresponding operations. Then, three further concepts are introduced: simple equivalence of imsets (with a related concept of normalized imset), the natural extension of an imset and finally the settled extension of an imset.

DEFINITION 1.9 (imset, multiset)
Every integer-valued function on $\mathcal{U}$ is called an *imset* (on $\mathcal{U}$). Recall that the class of all imsets is denoted by $Z(\mathcal{U})$. Nonnegative imsets (i.e. elements of $Z^+(\mathcal{U})$) are called *multisets* (*on* $\mathcal{U}$). Define basic operations with imsets like summing, subtracting, multiplying by integers coordinatewise. Trivial examples of imsets are *zero imset* (assigning zero to each set from $\mathcal{U}$) denoted by 0 and the indicator δ_T defined for every $T \in \mathcal{U}$.

Terminological remark I would like to explain the chosen terminology. The term 'multiset' is taken from [Aigner, 1979]. Looking for a proper designation of those integer-valued functions I decided to use the abbreviation *imset* (**i**nteger-valued **multiset**).

DEFINITION 1.10 (simple equivalence, normalized imset)
Having $u, v \in Z(\mathcal{U})$ define:

$$u \sim v \Leftrightarrow \exists k, l \in \mathbb{N} \quad k \cdot u = l \cdot v.$$

Evidently, $\sim$ is an equivalence on $Z(\mathcal{U})$. It will be called the *simple equivalence*. 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. The class of normalized imsets will be denoted by $Z_{norm}(\mathcal{U})$.

Notice that $0 \notin Z_{norm}(\mathcal{U})$. The following statement is a simple consequence of these definitions.

LEMMA 1.1 For each nonzero $u \in Z(\mathcal{U})$ there exists exactly one $v \in Z_{norm}(\mathcal{U})$ with $u \sim v$. The corresponding simple equivalence class has the form $\{k \cdot v; k \in \mathbb{N}\}$.

Sometimes (for example in [Studený, 1993b]), an alternative view on imsets is convenient: to consider them as integer-valued functions on exp N. Therefore we are going to introduce their natural extensions. We need the following:

LEMMA 1.2 Every imset $u \in Z(\mathcal{U})$ has a uniquely determined extension

$\bar{u} : \exp N \rightarrow \mathbb{Z}$ satisfying the following two conditions:

$$\Sigma\{\bar{u}(S); S \subseteq N\} = 0 \tag{N.1}$$

$$\forall r \in N \quad \Sigma\{\bar{u}(S); r \in S \subseteq N\} = 0 \tag{N.2}$$

This defines a one-to-one correspondence between $Z(\mathcal{U})$ and the class of integer-valued functions on exp N satisfying (N.1) and (N.2).

DEFINITION 1.11 (natural extension)
For any $u \in Z(\mathcal{U})$ the uniquely determined extension $\bar{u} : \exp N \rightarrow \mathbb{Z}$ satisfying (N.1) and (N.2) is called the *natural extension of u*.
Another type of extension will be also used for elements of $R(\mathcal{U})$ in the other parts.

DEFINITION 1.12 (settled extension)
A function $\underline{r} : \exp N \rightarrow \mathbb{R}$ is called *settled* iff $\underline{r}(S) = 0$ whenever card $S \leq 1$. For any $r \in R(\mathcal{U})$ its *settled extension* $\underline{r}$ is defined by that condition.

1.3 QUASIORDERINGS ON IMSETS

In this section special types of quasiorderings on imsets are studied: namely quasiorderings induced by subclasses of $R(\mathcal{U})$ by means of scalar product. We study their properties and find the conditions completely characterizing these orderings. Then the question of finding the largest subclass of $R(\mathcal{U})$ inducing a given quasiordering is answered.

A lot of care is devoted to orderings which can be induced by a finite subclass of $Z(\mathcal{U})$. It appears that they can be alternatively defined by "prescribing" a finite set of "positive" imsets which establishes all other "positive" imsets. The theory developed in this series of papers can be successively applied especially for these orderings. Therefore both ways of their determination are studied:

(a) by means of an **inducing class** (leading to the concept of *skeleton*)

(b) by means of an **establishing set** (leading to the concept of *base*).

1.3.1 Inducing of Quasiorderings

DEFINITION 1.13 (quasiordering, ordering, exhaustive class)

a) A binary relation $\leq$ on imsets is called a *quasiordering* iff it satisfies the following two conditions:
$u \leq u$ whenever $u \in Z(\mathcal{U})$ *reflexivity*
$[w \leq v \ \& \ v \leq u] \Rightarrow w \leq u$ whenever $u, v, w \in Z(\mathcal{U})$. *transitivity*
In case that it also satisfies the condition:
$[u \leq v \ \& \ v \leq u] \Rightarrow v = u$ whenever $u, v \in Z(\mathcal{U})$ *antisymmetry*
it is called an *ordering*.

b) Given $r, u \in R(\mathcal{U})$ define their *scalar product* as follows:
$\langle r, u\rangle = \Sigma_{S \in \mathcal{U}} r(S) \cdot u(S)$.

c) Given a class $C \subseteq R(\mathcal{U})$ define the binary relation $\lhd_C$ for imsets:
$v \lhd_C u \Leftrightarrow [\forall r \in C \quad \langle r, v\rangle \leq \langle r, u\rangle]$ whenever $u, v \in Z(\mathcal{U})$.
It is clearly a quasiordering, called the *scalar product quasiordering induced by C*.

d) A class $C \subseteq R(\mathcal{U})$ is called *exhaustive* iff it satisfies:
$[\langle r, u\rangle = 0$ for all $r \in C] \Rightarrow u = 0$ whenever $u \in Z(\mathcal{U})$.

The last condition is evidently equivalent to the antisymmetry of $\lhd_C$. Therefore in this case the relation $\lhd_C$ is called the *scalar product ordering induced by C*.

The essential properties of scalar product quasiorderings are formulated below. In fact, they are characterized as continuous linear quasiorderings.

ASSERTION 1.1 A binary relation $\leq$ on imsets is the scalar product quasiordering induced by some class $C \subseteq R(\mathcal{U})$ iff it satisfies the following three conditions:

$$u \leq u \quad \text{whenever } u \in Z(\mathcal{U}) \tag{V.1}$$

$$[v_1 \leq u_1 \ \& \ v_2 \leq u_2] \Rightarrow v_1 + v_2 \leq u_1 + v_2 \quad \text{whenever } u_1, u_2, v_1, v_2 \in Z(\mathcal{U}) \tag{V.2}$$

$$[\{\gamma_k\} \subseteq \mathbb{Q}^+ \ 0 \leq u_k \in Z(\mathcal{U}) \lim_{k\to\infty} \gamma_k \cdot u_k = u \in Z(\mathcal{U})] \Rightarrow 0 \leq u. \tag{V.3}$$

These conditions also imply:

$$v \leq u \Leftrightarrow v + w \leq u + w \quad \text{whenever } u, v, w \in Z(\mathcal{U}) \tag{V.4}$$

$$v \leq u \Leftrightarrow n \cdot v \leq n \cdot u \quad \text{whenever } u, v \in Z(\mathcal{U}),\ n \in \mathbb{N}. \tag{V.5}$$

Proof The necessity of (V.1)–(V.3) follows easily from elementary properties of scalar product (linearity and continuity). Conversely, supposing (V.1)–(V.3) put

$L = \{u \in Z(\mathcal{U}); 0 \preceq u\}$. Evidently, $v \preceq u$ iff $u - v \in L$ and L satisfies the following conditions:

$$u, v \in L \Rightarrow u + v \in L$$

$$[u_k \in L \quad u \in Z(\mathcal{U}) \quad \beta_k, \beta \in \mathbb{N} \quad \beta_k^{-1} \cdot u_k \to \beta^{-1} \cdot u] \Rightarrow u \in L$$

$$0 \in L.$$

Hence, by Lemma 9 in [Studený, 1993a] there exists a nonempty closed cone $K \subseteq R(\mathcal{U})$ with $L = K \cap Z(\mathcal{U})$. By Consequence 1 in [Studený, 1993a] $K = \{y \in R(\mathcal{U}); \forall r \in C \ \langle r, y\rangle \geq 0\}$ for some $C \subseteq R(\mathcal{U})$.

To derive (V.4) use (V.1) to get $w \preceq w$ and $-w \preceq -w$, then apply (V.2). The implication $\Rightarrow$ in (V.5) follows from (V.2). For $\Leftarrow$ put $\gamma_k = n^{-1}$, $u_k = n \cdot u - n \cdot v$ for each $k \in \mathbb{N}$. By (V.4) $0 \preceq u_k$ and using (V.3) derive $0 \preceq u - v$ i.e. $v \preceq u$ by (V.4). ■

1.3.2 Regular Cones

Of course, a scalar product ordering $\preceq$ can be induced by several classes. Let us ask whether a largest such class exists and what are its characteristics. It can be easily seen that a quasiordering $\lhd_C$ coincides with the quasiordering $\lhd_{\overline{con}(C)}$ where $\overline{con}(C)$ is the least closed cone[5] containing the class $C \subseteq R(\mathcal{U})$. Thus, a naive hypothesis might be the largest class inducing $\lhd_C$ is exactly $\overline{con}(C)$. Nevertheless, this hypothesis is false as the following example shows.

EXAMPLE 1.1 Suppose card $N \geq 3$, choose different $A, B \in \mathcal{U}$ and a positive irrational number γ. Put $C = \{r \in R(\mathcal{U}); r(A) \geq \gamma \cdot r(B)\}$. Evidently C is a closed cone. Considering $u \in Z(\mathcal{U})$ with $0 \lhd_C u$ it is easy to see that: $[u(K) = 0$, for $K \in \mathcal{U} \setminus \{A, B\}]$ & $[u(B) = -\gamma \cdot u(A)]$. Thus, $0 \lhd_C u$ iff $u = 0$ i.e. $0 \lhd_D u$ where $D = R(\mathcal{U})$!

The essence of the counterexample consists in the fact that the "boundary" of that cone "does not meet well" $Q(\mathcal{U})$. This motivates the following concept which allows us to answer our question.

DEFINITION 1.14 (regular cone)
A set $C \subseteq R(\mathcal{U})$ is called a *regular cone* iff it is a nonempty closed cone and $Q(\mathcal{U})$ is dense[6] in $C \cap (-C)$ where $(-C) = \{r \in R(\mathcal{U}); -r \in C\}$.

ASSERTION 1.2 Suppose that $\preceq$ is a scalar product quasiordering. Then the class $C_{\preceq} = \{r \in R(\mathcal{U}); \forall 0 \preceq u \in Z(\mathcal{U}) \ \langle r, u\rangle \geq 0\}$ is a regular cone inducing $\preceq$. Moreover, it is the largest class inducing $\preceq$ and the only regular cone inducing $\preceq$.

Proof Existence and uniqueness is proved in [Studený, 1993a] as Proposition 6a. Moreover, every class inducing $\preceq$ is evidently contained in $C_{\preceq}$. ■

---

[5]Remember that a set $C \subseteq R(\mathcal{U})$ is called a *cone* iff it is closed under summing and multiplying by nonnegative real numbers; C is *closed* iff it contains limits of all convergent sequences of its elements.

[6]A is *dense* in B iff every element of B is a limit of a sequence of elements of $A \cap B$.

Remark If C_1, $C_2 \subseteq R$ are two regular cones, then the quasiordering induced by C_1 is stronger than the quasiordering induced by C_2 iff $C_2 \subseteq C_1$ (see Proposition 6b in [Studený, 1993a]).

1.3.3 Establishing of Quasiorderings

Now, another possibility to define quasiorderings on imsets will be dealt with: namely by prescribing a certain set of positive imsets.

DEFINITION 1.15 (finitely established quasiordering)
A binary relation $\preceq$ on imsets is called a *finitely established linear quasiordering* iff there exists a nonempty finite set $E \subseteq Z(\mathcal{U})$ such that $\preceq$ is determined by the following property:

$$v \preceq u \Leftrightarrow \left[n \cdot (u - v) = \sum_{w \in E} k_w \cdot w \quad \text{for some } n \in \mathbb{N},\ k_w \in \mathbb{Z}^+ \right].$$

Note that this equality always defines a binary relation satisfying (V.1)–(V.2) i.e. a linear quasiordering. In that case we say that *E establishes* $\preceq$.

If $\preceq$ moreover satisfies the antisymmetry condition it is called a *finitely established linear ordering*.

A finitely established quasiordering is a special case of a scalar product quasiordering (see [Studený, 1993a], Proposition 7a). Nevertheless, the most interesting result can be achieved in the case of orderings:

ASSERTION 1.3 The following three conditions on a binary relation $\preceq$ on imsets are equivalent:

(a) $\preceq$ is a finitely established ordering

(b) $\preceq$ is a quasiordering established by a nonempty finite $G \subseteq Z(\mathcal{U})$ satisfying $[\exists q \in R(\mathcal{U})\ \forall u \in G\setminus\{0\}\ \langle q, u\rangle > 0]$.

(c) $\preceq$ is a scalar product ordering induced by a finite exhaustive class $C \subseteq Z(\mathcal{U})$.

Whenever any of these conditions is fulfilled there exists the least set of normalized imsets establishing $\preceq$ (necessarily finite).

Proof Use Proposition 8 in [Studený, 1993a]. The condition (a) corresponds to (iii) there, (b) to (ii) and (c) to (iv). ■

1.3.4 Base and Skeleton

The previous result naturally leads to the following concept.

DEFINITION 1.16 (base)
Supposing $\preceq$ is a finitely established ordering, the least establishing set of normalized imsets will be called the *base* of $\preceq$ and its elements *basic imsets for* $\preceq$.

Having a finitely established ordering by Assertion 1.3 we easily derive that there exists a finite subclass of $Z_{norm}(\mathcal{U})$ inducing it. Let us ask whether there exists the least such class. The answer is no, as the following example illustrates.

EXAMPLE 1.2 Suppose card $N \geq 3$, choose $A \in \mathcal{U}$ and put $E = \{\delta_B;\ B \in \mathcal{U}\backslash\{A\}\}$. It establishes the quasiordering $\leq$ described as follows:
$0 \leq u \Leftrightarrow [u(A) = 0\ \&\ \forall B \in \mathcal{U}\backslash\{A\}\ u(B) \geq 0]$ whenever $u \in Z(\mathcal{U})$.
It is evidently an ordering and the following classes of imsets
$K_1 = \{\delta_A, -\delta_A\} \cup \{\delta_B;\ B \in \mathcal{U}\backslash\{A\}\}$
$K_2 = \{\delta_A, -\delta_A\} \cup \{\delta_A + \delta_B;\ B \in \mathcal{U}\backslash\{A\}\}$
are both minimal subclasses of $Z_{norm}(\mathcal{U})$ inducing $\leq$.

Nevertheless, the existence of the least finite inducing subclass of $Z_{norm}(\mathcal{U})$ can be ensured as follows.

ASSERTION 1.4

(a) Let $\leq$ be a quasiordering established by a finite exhaustive set $E \subseteq Z(\mathcal{U})$. Then there exists the least finite subclass $A \subseteq Z_{norm}(\mathcal{U})$ inducing $\leq$.

(b) Moreover $\forall r \in A\ \exists u \in Z(\mathcal{U})\quad \langle r, u\rangle = 0\ \&\ \forall s \in A\backslash\{r\}\ \langle s, u\rangle > 0$.

Proof Use Proposition 7b from [Studený, 1993a].

Remark (warning) The previous assertion does not give the least subclass of $Z_{norm}(\mathcal{U})$ inducing the quasiordering! Another *infinite* inducing subclass may exist. For example, the natural ordering $\leq$ is generated by an exhaustive set $E = \{\delta_S;\ S \in \mathcal{U}\}$. The same class induces $\leq$. Nevertheless, in case card $N' \geq 3$ choose different $A, B \in \mathcal{U}$ and put: $K_{A,B} = \{n \cdot \delta_A + \delta_B;\ n \in \mathbb{N}\} \cup \{\delta_S;\ S \in \mathcal{U}\backslash\{A\}\}$. This class is an infinite subclass of $Z_{norm}(\mathcal{U})$ inducing the natural ordering which does not contain E.

The previous assertion motivates the following definition which concludes the section.

DEFINITION 1.17 (skeleton)
Supposing that $\leq$ is a quasiordering established by a finite exhaustive set of imsets, the least finite class of normalized imsets inducing $\leq$ will be called the *skeleton* of $\leq$ and its elements *skeletal imsets for* $\leq$.

REFERENCES

Aigner, M. [1979], *Combinatorial Theory*. Springer-Verlag, Berlin—Heidelberg—New York.

Brøndsted, A. [1983], *An Introduction to Convex Polytopes*. Springer-Verlag, Berlin—Heidelberg—New York—Tokyo.

Darroch, J. N., S. L. Lauritzen and T. P. Speed [1980], "Markov field theory and loglinear interaction models for contingency tables." *Annals of Statistics*, **8**(3), pp. 522–539.

Dawid, A. P. [1979], "Conditional independence in statistical theory." *Journal of the Royal Statistical Society B*, **41**(1), pp. 1–31.

Geiger, D. and J. Pearl [1990], "Logical and algorithmic properties of independence and their application to Bayesian networks." *Annals of Mathematics and Artificial Intelligence*, **2**(1/4), pp. 165–178.

Geiger, D. and J. Pearl [1993], "Logical and algorithmic properties of independence and graphical models." *Annals of Statistics*, **21**(4), pp. 2001–2021.

Howard, R. A. and J. E. Matheson [1981], "Influence diagrams." Chapter 8 in *The Principles and Applications of Decision Analysis 2*, Strategic Decision Group, Menlo Park, California.

Kiiveri, H., T. P. Speed and J. B. Carlin [1984], "Recursive causal model." *Journal of Australian Mathematical Society A*, **36**(1), pp. 30–52.

Lauritzen, S. L., T. P. Speed and K. Vijayan [1984], "Decomposable graphs and hypergraphs." *Journal of Australian Mathematical Society A*, **36**(1), pp. 12–29.

Matúš, F. [1992], "Ascending and descending conditional independence relations." In: *Information Theory Statistical Decision Functions Random Processes: transactions of the 11th Prague Conference* *vol. B*, edited by S. Kubik and J. A. Víšek, Kluwer, Dordrecht (also Academia, Prague), pp. 189–200.

Matúš, F. [1995], "Stochastic independence, algebraic independence and abstract connectedness." To appear in *Theoretical Computer Science A*.

Mouchart, M. and J.-M. Rolin [1984], "A note on conditional independence with statistical application." *Statistica*, **44**(4), pp. 557–584.

Moussouris, J. [1974], "Gibbs and Markov random systems with constraints." *Journal of Statistical Physics*, **10**(1), pp. 11–33.

NETWORKS [1990], "Special Issue on Influence Diagrams." *Networks*, **20**(5), pp. 486–700.

Pearl, J. [1988], *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, San Mateo, California.

Pearl, J., D. Geiger and T. Verma [1990], "The logic of influence diagrams." Chapter 4 in *Influence Diagrams, Belief Nets and Decision Analysis*, edited by R. M. Oliver and J. Q. Smith, John Wiley, Chichester.

Pearl, J. and A. Paz [1987], "Graphoids: graph-based logic for reasoning about relevance relations." In: *Advances in Artificial Intelligence—II*, edited by B. Du Boulay, D. Hogg and L. Steels, North-Holland, Amsterdam, pp. 357–363.

van Putten, C. and J. H. van Shuppen [1985], "Invariance properties of conditional independence relation." *Annals of Probability*, **13**(3), pp. 934–945.

Shachter, R. D. [1986], "Evaluating influence diagrams." *Operation Research*, **34**(6), pp. 871–882.

Smith, J. Q. [1989], "Influence diagrams for statistical modelling." *Annals of Statistics*, **17**(2), pp. 654–672.

Spohn, W. [1980], "Stochastic independence, causal independence and shieldability." *Journal of Philosophical Logic*, **9**(1), pp. 73–99.

Studený, M. [1989a], "Multiinformation and the problem of characterization of conditional independence relations." *Problems of Control and Information Theory*, **18**(1), pp. 3–16.

Studený, M. [1989b], "Attempts at axiomatic description of conditional independence." *Kybernetika*, **25**(suppl. 1/3), pp. 72–79.

Studený, M. [1992], "Conditional independence relations have no finite complete characterization." In: *Information Theory Statistical Decision Functions Random Processes: transactions of the 11th Prague Conference vol. B*, edited by S. Kubík and J. A. Víšek, Kluwer, Dordrecht (also Academia, Prague), pp. 377–396.

Studený, M. [1993a], "Convex cones in finite-dimensional real vector spaces." *Kybernetika*, **29**(2), pp. 180–200.

Studený, M. [1993b], "Description of conditional independence structures by means of imsets: a connection with product formula validity." In: *Uncertainty in Intelligent Systems*, edited by B. Bouchon-Meunier, L. Valverde and R. R. Yager, Elsevier, Amsterdam, pp. 179–194.

Ur, S. and A. Paz [1994], "The representation power of probabilistic knowledge by undirected graphs and directed acyclic graphs: a comparison." *International Journal of General Systems*, **22**(2), pp. 219–231.

Wermuth, N. and S. L. Lauritzen [1983], "Graphical and recursive models for contingency tables." *Biometrika*, **70**(3), pp. 537–552.

Wright, S. [1934], "The method of path coefficients." *The Annals of Mathematical Statistics*, **5**(3), pp. 161–215.

For a photo and biography of the author, please see 22(2), page 217, of this journal. New e-mail address of the author is studeny@utia.cas.cz.