Int. J. General Systems, Vol. 23, pp. 201–219
Reprints available directly from the publisher
Photocopying permitted by license only

DESCRIPTION OF STRUCTURES OF STOCHASTIC CONDITIONAL INDEPENDENCE BY MEANS OF FACES AND IMSETS
2nd part: basic theory[1]

MILAN STUDENÝ

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

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

Local Abstract (2nd part) The central concept of face (with respect to an arbitrary continuous linear ordering on imsets) and the corresponding deductive mechanism (facial implication) are introduced. The class of all faces constitutes a lattice which, in the case of finitely established orderings, is shown to be finite. Moreover, its atoms and co-atoms are characterized and two possible representations of faces are treated: by means of generating imsets and by means of portraits. Then a concrete ordering, called the structural ordering, is studied. Faces with respect to this ordering are identified with a class of dependency models including all models of probabilistic conditional independence structures.

INDEX TERMS:[2] Conditional independence, imset, face, structural ordering, structural semigraphoid.

PREFACE

This paper is the second installment of the work *Description of structures of stochastic conditional independence by means of faces and imsets*, a series of three papers. The purpose of the work as a whole is to present a new approach to description of probabilistic CI-structures (= conditional independence structures) and to relate it to classical methods of their description. This part contains the mathematical fundamentals of the theory. Before reading this paper the reader should be familiar with the concepts introduced in the first installment, subtitled *1st part: introduction and basic concepts*. The motivation of the theory developed here is explained and a global view on its construction is given there. Note that the last installment, subtitled *3rd part: examples of use and appendices*, consists of several examples of use of the theory and Conclusions where advantages and disadvantages of and prospects for the presented approach are discussed.

Note that each definition or result (throughout all series) is denominated by two numbers: the first one indicates the part where it can be found and the second one is its location within that part.

---

[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.

Let us outline the subject of this second installment in detail (concepts from the first installment are assumed to be known). The paper is divided into three sections; every section has an introductory paragraph describing its contents. The first section introduces the concept of *face* with respect to an arbitrary scalar product ordering of imsets. In the case of a finitely established linear ordering, subminimal and submaximal faces are characterized, two possible characterization of faces are derived, and the corresponding deductive mechanism, called *facial implication*, is described in terms of the skeleton. Note that the last result is essential for computer implementation of facial implication.[3] The second section studies a concrete example of such an ordering, called the *structural ordering*. It is characterized in two ways, namely by means of an establishing set of imsets and by means of an inducing class. The third section relates faces with respect to the structural ordering to certain dependency models, namely so-called *structural semigraphoids*; this identification makes it possible to use faces as mathematical tools for description of CI-structures. Specifically, using some information-theoretical concepts a structural face describing the corresponding CI-structure is assigned to every probability measure and therefore every probabilistic CI-structure is described by a structural semigraphoid. Moreover, the pertinent facial implication is shown to entail the probabilistic implication of CI-statements. Thus, semigraphoid derivability can be replaced by a more fitting approximation of the probabilistic implication, which is finitely implementable from a theoretical point of view, although it subsumes infinitely many inference rules of the semigraphoid type.

NOTATION

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 $\mathcal{U}$:

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

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

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

The class of all probability measures over N (see Def 1.1, §1.1.1) will be denoted by $\mathcal{P}(N)$, the simple equivalence (see Def 1.10, section 1.2) by $\sim$. Having disjoint sets $A, B \subseteq N$ the juxtaposition AB will stand for their union $A \cup B$.

Further, the set of real numbers will be denoted by $\mathbb{R}$, the set of nonnegative integers (including zero) by $\mathbb{Z}^+$ and the set of positive integers (natural numbers) by $\mathbb{N}$.

---

[3]The details are explained in the third installment.

Finally, the following symbols for classes of functions on $\mathcal{U}$ will be used:

$R(\mathcal{U})$	the class of real functions on $\mathcal{U}$
$Z(\mathcal{U})$	the class of integer-valued functions on $\mathcal{U}$
$Z_{norm}(\mathcal{U})$	the class of normalized integer-valued functions on $\mathcal{U}$, see Def 1.10
$Z^{+}(\mathcal{U})$	the class of nonnegative integer-valued functions on $\mathcal{U}$.

2.1 FACES

The central concept of this section is the concept of face introduced for every scalar product ordering $\preceq$ on imsets here. In section 2.3, just faces with respect to a so-called structural ordering will be used to describe CI-structures. In addition to the primary definition several further concepts are introduced: the face generated by a set of imsets, *facial implication* of imsets and *facial equivalence* on imsets.

Then, for the case that the ordering $\preceq$ is finitely established, some basic results are proved:

— every face is uniquely determined by its intersection with the base

— there exist finitely many faces

— every face is generated by a single imset (called the *generating imset*)

— faces are characterized by means of the skeleton (and therefore facial implication and equivalence are formalizable).

The last result leads to the concept of *portrait* facilitating computer representation of faces. Moreover, as consequences of preceding results characterizations of subminimal and submaximal faces are given. Subminimal faces (i.e. *atoms* of the lattice of faces) are shown to be just faces generated by basic imsets, while submaximal faces (i.e. *co-atoms* of the lattice of faces) correspond uniquely to skeletal imsets.

Throughout this section a scalar product ordering $\preceq$ on imsets (see Def 1.13, §1.3.1) will be dealt with. The set of "positive" imsets (with respect to $\preceq$) will be essential in the sequel:

Notation $Z(\preceq) = \{u \in Z(\mathcal{U});\quad 0 \preceq u\}$.

2.1.1 *Definition of Face*

DEFINITION 2.1 (face)
A set $F \subseteq Z(\preceq)$ is called a *face* iff it satisfies the following three conditions:

$0 \in F$	*nontriviality*	(F.0)
$u, v \in F \Rightarrow u + v \in F$	*composition*	(F.1)
$u, v \in Z(\preceq)\ u + v \in F \Rightarrow u, v \in F$	*decomposition*	(F.2)

Terminological remark This terminology is motivated by an analogy with the theory of convex polytopes [Brøndsted, 1983] where the concept of face has a central role. I think it is possible to prove that every subset of $Z(\preceq)$ satisfying (F.0)-(F.2) is the intersection of $Z(\mathcal{U})$ with a zero-containing face (in Brøndsted's sense) of the cone determined by $\leq$.

Trivial examples of faces are $\{0\}$ and $Z(\leq)$ i.e. the least and the largest face. Evidently, the intersection of every nonempty collection of faces is a face too. Therefore it holds:

FACT 2.1 For any $L \subseteq Z(\leq)$ there exists the least face containing L.

It is now easy to see that the class of faces constitutes a complete lattice[4] with respect to inclusion. Moreover, Fact 2.1 makes the following definition possible.

DEFINITION 2.2 (generated face, generating imset)
For any $L \subseteq Z(\leq)$, the least face containing L will be called the *face generated by* L and denoted by $\bar{f}(L)$. Whenever a face F has the form $\bar{f}(\{u\})$ where $u \in Z(\leq)$ say that u is the *generating imset of* F.

In fact, every face generated by a finite set is generated by a single imset:

LEMMA 2.1 For any finite $L \subseteq Z(\leq), \bar{f}(L) = \bar{f}(\{\Sigma_{u\in L} u\})$.

Proof Let $F = \bar{f}(\{\Sigma_{u\in L} u\})$. For each $v \in L$ write $\Sigma_{u\in L} u = v + (\Sigma_{w\in L\setminus\{v\}} w)$ and by decomposition derive $v \in F$. Thus, $L \subseteq F$ gives $\bar{f}(L) \subseteq F$. Conversely, by composition $(\Sigma_{u\in L} u) \in \bar{f}(L)$ and hence $F \subseteq \bar{f}(L)$. ■

2.1.2 Facial Implication

Faces introduce a certain deductive mechanism for imsets, more precisely for elements of $Z(\leq)$. It will be shown that composition and decomposition may be interpreted as inference rules for faces. (Nontriviality plays the role of an axiom.)

DEFINITION 2.3 (facial implication, facial quasiordering)
Suppose that $L \subseteq Z(\leq)$ and $u \in Z(\leq)$. Say that L *facially implies* u and write $L \mapsto u$ iff $L \subseteq F \Rightarrow u \in F$ whenever F is a face.
Whenever $u, v \in Z(\leq)$ write also $u \mapsto v$ instead of $\{u\} \mapsto v$. This binary relation on $Z(\leq)$ is clearly a quasiordering called the *facial quasiordering*.

LEMMA 2.2 Let $L \subseteq Z(\leq)$ and $u \in Z(\leq)$. Then $L \mapsto u$ iff any of the following three conditions hold:

(a) $u \in \bar{f}(L)$

(b) there exists a finite subset $L' \subseteq L$ and numbers $k_v \in \mathbb{Z}^+$ $(v \in L')$ such that $(\Sigma_{v\in L'} k_v \cdot v - u) \in Z(\leq)$, where we accept the convention $\Sigma_{v\in\emptyset} k_v \cdot v = 0$

(c) u is derivable from L by means of (F.0)–(F.2) i.e. there exists a derivation sequence $w_1, \ldots, w_n$ where $w_n = u$ such that for each w_j either $w_j \in L$ or $w_j = 0$ (i.e. (F.0) is used) or w_j is a direct consequence of some preceding w_is by virtue of (F.1) or (F.2).

[4]A *complete lattice* is a partially ordered set every subset A of which has a *supremum* (i.e. the least element greater than all elements of A) denoted by sup A and an *infimum* (i.e. the greatest element less than all elements of A) denoted by inf A.

Proof I. $[L \mapsto u] \Rightarrow$ (a) is evident from Def 2.2.

II. (a) $\Rightarrow$ (b) Put $F = \{u \in Z(\leqslant);\ u$ satisfies the condition from (b)$\}$. Evidently $L \subseteq F$ and it is easy to verify that F is a face.

III. (b) $\Rightarrow$ (c) The derivation sequence can be described as follows:

- first list the elements of L'
- imsets $k_v \cdot v$ ($v \in L'$) are derived either using (F.0) (the case $k_v = 0$) or by a consecutive application of (F.1) (the case $k_v > 0$)
- $\Sigma_{v \in L'} k_v \cdot v$ is derived by virtue of (F.1)
- $\Sigma_{v \in L'} k_v \cdot v = u + (\Sigma_{v \in L'} k_v \cdot v - u)$ makes it possible to derive u by virtue of (F.2).

IV. (c) $\Rightarrow [L \mapsto u]$. By induction every member of the derivation sequence belongs to any face containing L. ■

It is easy to see by Lemma 2.2(b) that the facial quasiordering is linear:

FACT 2.2 $[u_1 \mapsto v_1\ \&\ u_2 \mapsto v_2] \Rightarrow u_1 + u_2 \mapsto v_1 + v_2$
whenever $u_1, u_2, v_1, v_2 \in Z(\leq)$.

2.1.3 Facial Equivalence

As $\mapsto$ is not an ordering on $Z(\leq)$, a further natural step is to introduce an equivalence remodelling it in an ordering.

DEFINITION 2.4 (facial equivalence)
For $u, v \in Z(\leq)$ say that u is *facially equivalent to* v and write $u \approx v$ iff

$$u \in F \Leftrightarrow v \in F \quad \text{whenever } F \text{ is a face.}$$

The corresponding factor space[5] will be denoted by $Z(\leq)/_{\approx}$.

Of course, $u \approx v$ iff $[u \mapsto v\ \&\ v \mapsto u]$. Therefore, we get from Lemma 2.2(a):

FACT 2.3 $u \approx v$ iff $\mathcal{F}(\{u\}) = \mathcal{F}(\{v\})$ whenever $u, v \in Z(\leq)$.

Let us compare facial equivalence with simple equivalence (see Def 1.10, section 1.2). By (V.5) (see Assertion 1.1, §1.3.1) the simple equivalence $\sim$ respects $Z(\leq)$ (i.e. for $u, v \in Z(\mathcal{U})$ $u \sim v \in Z(\leq) \Rightarrow u \in Z(\leq)$) and therefore $\sim$ can be considered on $Z(\leq)$. Moreover, $\sim$ similarly respects an arbitrary face. Hence,

FACT 2.4 $u \sim v \Rightarrow u \approx v$ whenever $u, v \in Z(\leq)$.

As $\mapsto$ is conformable to $\approx$-equivalence classes, it can be considered as a binary operation on the factor space $Z(\leq)/_{\approx}$. An analogous conclusion holds for $+$ (use Fact 2.2). Altogether:

SUMMARY 2.1 The factor space $Z(\leq)/_{\approx}$ is endowed with two operations:

---

[5]The factor space is the set of equivalence classes.

a) the ordering $\mapsto$ (interpret $u \mapsto v$ as "u is greater than v") b) $+$.
They are associated by the following property:

$$[\tilde{u}_1 \mapsto \tilde{v}_1 \quad \& \quad \tilde{u}_2 \mapsto \tilde{v}_2] \Rightarrow \tilde{u}_1 + \tilde{u}_2 \mapsto \tilde{v}_1 + \tilde{v}_2 \quad \text{whenever } \tilde{u}_1, \tilde{u}_2, \tilde{v}_1, \tilde{v}_2 \in Z(\leq)/_{\sim}.$$

Hence, one can easily derive that $\sup(\tilde{u}, \tilde{v}) = \tilde{u} + \tilde{v}$ whenever $\tilde{u}, \tilde{v} \in Z(\leq)/_{\sim}$.

Note that $Z(\leq)/_{\sim}$ is in fact isomorphic to the lattice of faces (see Lemma 2.5 in §2.1.6).

2.1.4 Faces for Finitely Established Orderings

In the rest of this section $\preccurlyeq$ will be considered to be a finitely established linear ordering (Def 1.15, §1.3.3). The first intention is to represent faces in its base (see Def 1.16, §1.3.4).

LEMMA 2.3 For a finitely established $\leq$, consider a finite set $T \subseteq Z(\leq)$ establishing it. Then it holds: $F_1 \cap T \subseteq F_2 \cap T \Rightarrow F_1 \subseteq F_2$ whenever F_1, F_2 are faces.

Proof Take $u \in F_1$ and write $n \cdot u = \Sigma_{v \in T} k_v \cdot v$ ($n \in \mathbb{N}$, $k_v \in \mathbb{Z}^+$). Whenever $k_v > 0$ then $v \in F_1 \cap T$ (decomposition for F_1) and therefore $v \in F_2 \cap T$. By composition for F_2 derive then $u \in F_2$. ■

LEMMA 2.4 Suppose that $\leq$ is finitely established and let $E \subseteq Z(\leq)$ be its base. The following conditions for $0 \neq u \in Z(\leq)$ are equivalent:

(a) $u \in E$

(b) $u \in Z_{norm}(\mathcal{U})$ & $[\forall 0 \neq v \in Z(\leq) \quad u \mapsto v \Rightarrow u \sim v]$

(c) $\{n \cdot u;\ n \in \mathbb{Z}^+\}$ is a face.

Proof (a) $\Rightarrow$ (b) Consider $0 \neq v \in Z(\leq)$ with $u \mapsto v$, by Lemma 2.2(b) $n \cdot u - v \in Z(\leq)$ for some $n \in \mathbb{Z}^+$, as $v \neq 0$ necessarily $n \in \mathbb{N}$. As E establishes $\leq$ write $m \cdot (n \cdot u - v) = \Sigma_{w \in E} k_w \cdot w$ and $p \cdot v = \Sigma_{w \in E} l_w \cdot w$ where $m, p \in \mathbb{N}$, $k_w, l_w \in \mathbb{Z}^+$. Hence $(mpn - pk_u - ml_u) \cdot u = \Sigma_{w \in E \setminus \{u\}} (pk_w + ml_w) \cdot w$. Necessarily $mpn - pk_u - ml_u \leq 0$ (otherwise u is expressed by means of $E \setminus \{u\}$, therefore $E \setminus \{u\}$ establishes $\leq$ and it contradicts the definition of the base). Thus, the antisymmetry condition for $\leq$ implies that the expression above vanishes and hence $l_w = 0$ for $w \in E \setminus \{u\}$. Therefore $p \cdot v = l_u \cdot u$ gives easily $u \sim v$.

(b) $\Rightarrow$ (c) To verify (F.2) for $\{n \cdot u,\ n \in \mathbb{Z}^+\}$ write $n \cdot u = v + w$ where $v, w \in Z(\leq)$. In the nontrivial case $v \neq 0$ by Lemma 2.2(b) $u \mapsto v$ and using (b) $u \sim v$. Lemma 1.1, section 1.2 implies then $v = k \cdot u$ for $k \in \mathbb{N}$.

(c) $\Rightarrow$ (a) Consider two different faces: $\{n \cdot u,\ n \in \mathbb{Z}^+\}$ and $\{0\}$. As E establishes $\leq$ by Lemma 2.3 $\{n \cdot u;\ n \in \mathbb{Z}^+\} \cap E \neq \emptyset$ and hence by Lemma 1.1 $u \in E$. ■

DEFINITION 2.5 (distinguishing of faces)
A set $T \subseteq Z(\leq)$ *distinguishes faces* iff

$$F_1 \neq F_2 \Rightarrow F_1 \cap T \neq F_2 \cap T \quad \text{whenever } F_1, F_2 \text{ are faces.}$$

Now, the first theorem can be formulated.

THEOREM 2.1
Suppose that $\leq$ is a finitely established linear ordering on imsets. Then its base E is the least set of normalized imsets distinguishing faces (i.e. faces are determined uniquely by their intersection with the base). Moreover, every face F has the form $F = \bar{f}(F \cap E)$.

Proof By Lemma 2.3 E distinguishes faces. Suppose that $T \subseteq Z_{norm}(\mathcal{U}) \cap Z(\leq)$ distinguishes faces too. Consider $u \in E$ and put $F_1 = \{n \cdot u;\ n \in \mathbb{Z}^+\}$, $F_2 = \{0\}$. By Lemma 2.4 F_1 is a face, as T distinguishes faces $F_1 \cap T \neq F_2 \cap T = \emptyset$. Nevertheless, $T \subseteq Z_{norm}(\mathcal{U})$ implies by Lemma 1.1 $u \in T$ and the inclusion $E \subseteq T$ is verified. Finally consider an arbitrary face F and put $K = \bar{f}(F \cap E)$. Evidently $K \subseteq F$ and as $F \cap E \subseteq K \cap E$, $F \subseteq K$ by Lemma 2.3. ■

2.1.5 *Atomic Faces*

The preceding theorem has two important consequences.

CONSEQUENCE 2.1 Suppose that $\leq$ is finitely established. A face is an atom[6] of the lattice of faces iff it has the form $\{n \cdot u;\ n \in \mathbb{Z}^+\}$ for some basic imset u (for $\leq$).

Proof For $F = \{n \cdot u;\ n \in \mathbb{Z}^+\}$ where $u \in E$ (= the base) consider a face $K \subseteq F$. In the case that $K \neq \{0\}$ we can find $0 \neq v \in K$ and derive $u \sim v$. Hence $u \in K$ and therefore $F \subseteq K$.

Conversely, let F be an atom of the lattice of faces. By Theorem 2.1 $F \cap E \neq \emptyset$, so we choose $u \in F \cap E$ and put $K = \{n \cdot u;\ n \in \mathbb{Z}^+\}$. By Lemma 2.4(c) K is a face and $K \cap E \subseteq F \cap E$ implies by Lemma 2.3 $K \subseteq F$. As F is an atom derive $F = K$. ■

Remark It can be proved that $\leq$ is finitely established iff the set $\{u \in Z(\leq);\ \{n \cdot u;\ n \in \mathbb{Z}^+\}$ is a face$\}$ is finite and establishes $\leq$.

CONSEQUENCE 2.2 Supposing that $\leq$ is finitely established there exist finitely many faces.

Proof This follows easily from Theorem 2.1 as the base has finitely many subsets. ■

2.1.6 *Generating Imsets*

LEMMA 2.5 The following conditions are equivalent (for a scalar product ordering $\leq$):

(i) there exist finitely many faces
(ii) the factor space $Z(\leq)/_{\sim}$ is finite.

Whenever either of them holds:

[6] An element of a lattice is called an *atom* iff the only different less element than it is the least element of the lattice.

a) Each face has a generating imset.
b) The mapping $u \to \bar{f}(\{u\})$ (assigning faces to imsets) is constant on $\approx$-equivalence classes. It defines a one-to-one correspondence between $Z(\leq)/_{\approx}$ and the lattice of faces.

This correspondence is isotone[7] (i.e. isomorphism of lattices).

Proof By Fact 2.3 the mapping $u \to \bar{f}(\{u\})$ can be understood as an injective mapping of $Z(\leq)/_{\approx}$ into the class of faces. Therefore (i) $\Rightarrow$ (ii). Now, suppose (ii). By proving a) it will be shown that the mapping is onto. Hence (i) follows and the mapping is one-to-one. The simple argument $[u \mapsto v \text{ iff } \bar{f}(\{v\}) \subseteq \bar{f}(\{u\})]$ will say the rest. Thus, to show a) consider an arbitrary face F and note that every $\approx$-equivalence class intersecting F is contained in F. Therefore using (ii) conclude that F is the union of finitely many $\approx$-equivalence classes $L_1, \ldots, L_k$. Choose $u_i \in L_i$ and put u $\sum_{i=1}^{k} u_i$. As every element of F is equivalent with some u_i: get $F \subseteq \bar{f}(\{u_1, \ldots, u_k\}) \subseteq F$ and hence by Lemma 2.1 $F = \bar{f}(\{u\})$. ■

THEOREM 2.2
If $\leq$ is a finitely established linear ordering every face has a generating imset.

Proof Combine Consequence 2.2 and Lemma 2.5. ■

Remark Intuition suggests that the previous result holds for arbitrary scalar product ordering. However, I did not try to prove this as the result above is sufficient for the purposes of this work.

2.1.7 *Skeleton and Faces*

Finally, faces and corresponding concepts will be described by means of any finite inducing class, especially by means of the skeleton (see Def 1.17, §1.3.4) in case that it exists. Note that all these results can be achieved even if there exists a finite $C \subseteq R(\mathcal{U})$ inducing $\leq$.

ASSERTION 2.1 Suppose that $\leq$ is a finitely established ordering and $C \subseteq Z(\mathcal{U})$ is any finite class inducing it (its existence follows from Assertion 1.3, §1.3.3). Suppose $L \subseteq Z(\leq)$ is finite and $u \in Z(\leq)$. Then $L \mapsto u$ iff the following condition holds:

$$\forall r \in C \quad \langle r, u\rangle > 0 \Rightarrow [\langle r, v\rangle > 0 \text{ for some } v \in L].$$

Proof By Lemma 2.2(b) $[L \mapsto u]$ means that there exists $k_v \in \mathbb{Z}^+ (v \in L)$ such that $0 \leq. \sum_{v\in L} k_v \cdot v - u$ i.e. $\forall r \in C\ 0 \leq \langle r, \sum_{v\in L} k_v \cdot v - u\rangle = \sum_{v\in L} k_v\langle r, v\rangle - \langle r, u\rangle$. As C is finite, it is easy to see that the condition $[\exists k_v \in \mathbb{Z}^+ (v \in L)\ \forall r \in C\ \langle r, u\rangle \leq \sum_{v\in L} k_v\langle r, v\rangle]$ is equivalent to the desired condition. ■

CONSEQUENCE 2.3 Suppose that $\leq$ is finitely established, $C \subseteq Z(\mathcal{U})$ is a finite class inducing $\leq$; $u, v \in Z(\leq)$. Then $u \approx v$ iff $\forall r \in C\ \langle r, u\rangle > 0 \Leftrightarrow \langle r, v\rangle > 0$.

Proof $u \approx v \Leftrightarrow [u \mapsto v\ \&\ v \mapsto u]$; use Assertion 2.1. ■

---

[7]i.e. it preserves the ordering.

THEOREM 2.3
Suppose that $\leq$ is a finitely established linear ordering and $C \subseteq Z(\mathcal{U})$ is any finite class inducing $\leq$ (for example the skeleton in case that it exists). Then every face F has the form $F = \{u \in Z(\leq); \forall r \in D \ \langle r, u\rangle = 0\}$ where D is a subset of C. Conversely, any set of this form is a face.

Proof By Theorem 2.2 there exists $u \in Z(\leq)$ with $F = \bar{f}(\{u\})$ i.e. by Lemma 2.2(a) $F = \{v \in Z(\leq); u \mapsto v\}$. We put $D = \{r \in C; \langle r, u\rangle = 0\}$ and by Assertion 2.1 get $F = \{v \in Z(\leq); \forall r \in C \ \langle r, v\rangle > 0 \Rightarrow r \in C \setminus D\}$ and hence the desired form. The second statement can be shown by verifying (F.0)–(F.2). ■

2.1.8 Co-atomic Faces and Portrait

CONSEQUENCE 2.4 Suppose that $\leq$ is an ordering established by a finite exhaustive subset of $Z(\mathcal{U})$ and A is its skeleton. A face is a co-atom[8] of the lattice of faces iff it has the form $\{u \in Z(\leq); \langle r, u\rangle = 0\}$ for some $r \in A$.
Moreover, faces of this form are different for different $r \in A$.

Proof For card $N = 2$ the statement is trivial. Suppose card $N \subsetneq 3$.

I. $\forall r \in A \ \{u \in Z(\leq); \langle r, u\rangle = 0\} \neq Z(\leq)$.
As card $A \geq 2$ (otherwise $\leq$ is not an ordering) Assertion 1.4(b), §1.3.4 can be used.

II. F is a co-atom $\Rightarrow \exists r \in A \quad F = \{u \in Z(\leq); \langle r, u\rangle = 0\}$.
By Theorem 2.3 find $D \subseteq A$ with $F = \{u \in Z(\leq); \forall s \in S \ \langle s, u\rangle = 0\}$. Evidently $D \neq \emptyset$; choose $r \in D$ and put $K = \{u \in Z(\leq); \langle r, u\rangle = 0\}$. As K is a face containing F and by I. $K \neq Z(\leq)$ derive easily $F = K$.

III. $\forall r, s \in A \ \{u \in Z(\leq); \langle r, u\rangle = 0\} \subseteq \{u \in Z(\leq); \langle s, u\rangle = 0\} \Rightarrow r = s$.
In case $r \neq s$ use Assertion 1.4(b) to get the contradiction.

IV. $\forall r \in A \quad F = \{u \in Z(\leq); \langle r, u\rangle = 0\} \Rightarrow F$ is a co-atom.
Supposing K is a face with $F \subseteq K$ by Theorem 2.3 find $D \subseteq A$ with $K = \{u \in Z(\leq); \forall s \in D \ \langle s, u\rangle = 0\}$. By III. for each $s \in D$ derive $r = s$ i.e. $D \subseteq \{r\}$. In case $D = \{r\}$ get $K = F$, in case $D = \emptyset$ get $K = Z(\leq)$. ■

Remark This result does not hold if the skeleton is replaced by a minimal finite inducing class. Example 1.2 in §1.3.4 can provide a counterexample.

A useful concept allowing faces to be represented in a computer concludes this section.

DEFINITION 2.6 (portrait)
Suppose that $\leq$ is a finitely established ordering and $A \subseteq Z_{norm}(\mathcal{U})$ is a finite class inducing it (for example the skeleton). Whenever $u \in Z(\leq)$ the set $A_u = \{a \in A; \langle a, u\rangle > 0\}$ is called the *portrait of u in A*.
Similarly, for a face F (or simply a subset of $Z(\leq)$) its *portrait* is defined as the union of portraits of its elements: $A_F = \{a \in A; \langle a, v\rangle > 0$ for some $v \in F\}$.

[8]An element of a lattice is called a *co-atom* iff the only different greater element than it is the greatest element of the lattice.

It is easy to derive the following using Assertion 2.1 and Consequence 2.3:

SUMMARY 2.2 Portraits can describe both facial implication and equivalence:

a) $L \mapsto u$ iff $A_u \subseteq A_L$ whenever $L \subseteq Z(\leq)$, $u \in Z(\leq)$

b) $u \approx v$ iff $A_u = A_v$ whenever $u, v \in Z(\leq)$.

Thus, the portrait of a face is exactly the portrait of its generating imsets and portraits can be used for isomorphic description of faces:

c) $F_1 \subseteq F_2$ iff $A_{F1} \subseteq A_{F2}$ whenever F_1, F_2 are faces

or description of elements of $Z(\leq)/_{\approx}$ (see Lemma 2.5); owing to b) we can take $A_{\tilde{u}}$ as A_u for $u \in \tilde{u} \in Z(\leq)/_{\approx}$, by Summary 2.1, get:

d) $A_{\sup(\tilde{u},\tilde{v})} = A_{\tilde{u}+\tilde{v}} = A_{u+v} = A_u \cup A_v = A_{\tilde{u}} \cup A_{\tilde{v}}$.

2.2 STRUCTURAL IMSETS

In this section, a concrete finitely established linear ordering on imsets is introduced. It is called the *structural ordering* as faces with respect to this ordering serve as models of CI-structures in the next section. The ordering is defined by prescribing an establishing set. Nevertheless, an equivalent definition by means of an inducing class is incorporated: the largest inducing class is characterized. This makes it possible to derive several necessary conditions for structural imsets (i.e. imsets "positive" with respect to the structural ordering). Finally, the base is found and the question of finding the skeleton is discussed.

2.2.1 Structural Ordering

First, we introduce both possible establishing sets.

DEFINITION 2.7 (elementary and semielementary imsets)
An imset u is called *semielementary* iff its natural extension (see Def 1.11, section 1.2) has the form:

$$\bar{u} = \delta_{K \cup L} - \delta_K - \delta_L + \delta_{K \cap L} \quad \text{where } K, L \subseteq N.$$

The set of semielementary imsets will be denoted by E_{sem}. Moreover, an imset u is called *elementary* iff its natural extension has the form:

$$\bar{u} = \delta_{S \cup T} - \delta_S - \delta_T + \delta_{S \cap T} \quad \text{where } S, T \subseteq N \quad \text{card } S \setminus T = \text{card } T \setminus S = 1.$$

The set of elementary imsets will be denoted by E in the sequel.

LEMMA 2.6 The set of elementary imsets E is nonempty and finite, and
a) $\exists q \in R(\mathcal{U})\ \forall u \in E\ (q, u) > 0$
b) E is exhaustive (see Def 1.13d, §1.3.1).

Proof a) Define the sequence $\{r_k\}_{k\in\mathbb{N}}$ as follows:

$$r_1 = 0 \quad r_2 = 1 \quad r_{k+1} = 2\cdot r_k - r_{k-1} + 1 \quad \text{for } k \geq 2$$

and introduce $q \in Z^+(\mathcal{U})$ by the equality: $q(S) = r_{\mathrm{card}S}$ whenever $S \in \mathcal{U}$. It is easy to see that $\forall u \in E\ \langle q, u\rangle = 1$.

b) Consider $r \in Z(\mathcal{U})$ with $[\forall u \in E\ \langle r, u\rangle = 0]$. If $A \subset N$ and card $A = 2$ then $\delta_A \in E$ and therefore $r(A) = \langle r, \delta_A\rangle = 0$. Then by induction on card A prove $r(A) = 0$; it suffices to find a proper $u \in E$ of the form $\bar{u} = \delta_A - \delta_S - \delta_T + \delta_{S\cap T}$ with $A = S \cup T$. ■

Definition 2.8 (structural ordering, structural imset)
The *structural ordering* is the ordering established by the set of elementary imsets (use Assertion 1.3(b), §1.3.3). It will be denoted by $\prec$. An imset u is called *structural* iff $0 \prec u$. The class of structural imsets will be denoted by $Z(\prec)$.

Terminological remark Later, we will see that the natural identification of faces with dependency models is one-to-one for this ordering (see Lemma 2.9 in §2.3.1). Therefore such faces "precisely" correspond to structures described by dependency models. This specificity of the ordering motivated the terminology.

2.2.2 *Convex Set Functions*

Now, the class of completely convex functions, shown later to be the largest class inducing $\prec$, will be introduced.

Definition 2.9 (convex and completely convex set functions)
A set function m: exp $N \to \mathbb{R}$ is called *convex* iff it satisfies the condition of convexity: $m(K \cup L) + m(K \cap L) \geq m(K) + m(L)$ whenever $K, L \subseteq N$.
A set function $m \in R(\mathcal{U})$ (on $\mathcal{U}$ only!) is called a *completely convex set function* iff its settled extension $\underline{m}$ (see Def 1.12, section 1.2) is convex.
Finally, $m \in Z^+(\mathcal{U})$ is called *a convex multiset* iff it is a completely convex set function.

Terminological remark The terminology is taken from game theory (see [Rosenmüller and Weidner, 1974]). Some readers may prefer to call these functions *supermodular* (corresponding to superadditive functions in measure theory).

Lemma 2.7 a) For m: exp $N \to \mathbb{R}$ the following two conditions are equivalent:

(i) $m(K \cup L) + m(K \cap L) \geq m(K) + m(L)$ $K, L \subseteq N$

(ii) $m(S \cup T) + m(S \cap T) \geq m(S) + m(T)$

whenever, $S, T \subseteq N$ card $S \setminus T$ = card $T \setminus S = 1$.

b) Every convex settled function m : exp $N \to \mathbb{R}$ is nondecreasing and hence nonnegative.

Proof a) The implication $(i) \Rightarrow (ii)$ is trivial. Conversely, supposing (ii) prove the inequality in (i) by induction on t = card $((K \setminus L) \cup (L \setminus K))$. With $t > 2$ suppose card $K \setminus L \geq 2$ (otherwise replace K by L), choose $x \in K \setminus L$ and write:

$$m(K \cup L) + m(K \cap L) - m(K) - m(L) = [m(K \cup L) + m(K \setminus \{x\}) - m(K)$$
$$- m(K \cup L \setminus \{x\})] + [m(K \cup L \setminus \{x\}) + m(K \cap L) - m(L) - m(K \setminus \{x\})].$$

Both expressions in square brackets are nonnegative owing to the induction assumption.

b) Whenever $U \subseteq V \subseteq N$ with card $V \setminus U = 1$ the inequality $m(U) \leq m(V)$ follows from (i) where $K = U$, $L = V \setminus U$ (vanishing outside $\mathcal{U}$). ■

2.2.3 Characterization of Structural Imsets

The most important equivalent definition of the structural ordering is contained in the following theorem.

THEOREM 2.4

a) The structural ordering is established by the class of semielementary imsets.
b) The structural ordering is induced by the class of completely convex set functions, which is the largest class inducing it.[9]

Proof Supposing $u \in Z(\mathcal{U})$, it suffices to show that the following conditions are equivalent:

$$\exists n \in \mathbb{N} \quad k_v \in \mathbb{Z}^+ \; (v \in E) \quad n \cdot u = \sum_{v \in E} k_v \cdot v \tag{2.1}$$

$$\exists n \in \mathbb{N} \quad k_v \in \mathbb{Z}^+ \; (v \in E_{\mathrm{sem}}) \quad n \cdot u = \sum_{v \in E_{\mathrm{sem}}} k_v \cdot v \tag{2.2}$$

$$\langle m, u \rangle \geq 0 \text{ for every completely convex set function } m \in R(\mathcal{U}) \tag{2.3}$$

Since $E \subseteq E_{sem}$, (2.1) $\Rightarrow$ (2.2). Whenever $v \in E_{sem}$ and m is a completely convex function, then $\langle m, v \rangle \geq 0$. Therefore, supposing (2.2), the equality $n\langle m, u\rangle = \Sigma_{v \in E_{sem}} k_v \langle m, v\rangle \geq 0$ gives (2.3). Finally, to show (2.3) $\Rightarrow$ (2.1) use Lemma 10b in [Studený, 1993] (clearly by Lemma 2.7a the class of completely convex functions can be written as $\{m \in R(\mathcal{U}); \forall u \in E \; \langle m, u\rangle \geq 0\}$). Because of the class of completely convex set functions is a regular cone by Assertion 1.2, §1.3.2 it equals $C_{\prec}$. ■

REMARK 2.1 (how to recognize structural imsets)
Consider an imset u and the task of ascertaining whether u is structural. By definition, the direct method to prove that u is structural consists in "decomposition of a multiple of u into elementary imsets". Nevertheless, by Theorem 2.4a it suffices to decompose the multiple into semielementary imsets. This is sometimes easier as the class of semielementary imsets is wider. Moreover, the set of elementary (or semielementary) imsets can be reduced to $\bar{E} = \{v \in E; \bar{v} = \delta_{SUT} - \delta_S - \delta_T + \delta_{S \subset T}$ where $SUT \subseteq A \in \mathcal{U}$ with $u(A) \neq 0\}$. The proof is left to the reader as a simple exercise.
Note that in case card $N \leq 4$ one can decompose without first multiplying as both approaches are equivalent. This is shown in [Studený, 1991] §4, 5. Nevertheless, a more effective method is possible in case that the skeleton is at our disposal (see the discussion below Consequence 2.5, §2.2.4).

---

[9]Note that it can be shown that the class of convex multisets induces $\prec$ too.

Theorem 2.4b can serve as a tool to derive various necessary conditions for structural imsets. For example, whenever m is a convex multiset then

$$\langle m, u\rangle \geq 0 \quad \text{for each structural imset } u.$$

Concretely, convex multisets can be found as follows:

LEMMA 2.8 Suppose that $\mathcal{T} \subseteq \mathcal{U}$ is ascending, i.e. $S \in \mathcal{T}, S \subseteq T \in \mathcal{U} \Rightarrow T \in \mathcal{T}$. Denote the system of its minimal sets by $\mathcal{T}_{min}$ and introduce the multiset $m_{\mathcal{T}}$ (the *evaluation* of $\mathcal{T}$) by defining directly its settled extension $\underline{m_{\mathcal{T}}}$:

$$\underline{m_{\mathcal{T}}}(K) = 0 \quad \text{whenever } K \in (\exp N) \setminus \mathcal{T}$$

$$m_{\mathcal{T}}(K) = 1 \quad \text{whenever } K \in \mathcal{T}_{min},$$

and for $K \in \mathcal{T} \setminus \mathcal{T}_{min}$ define $m_{\mathcal{T}}(K)$ successively (by induction on card K):
$m_{\mathcal{T}}(K) = \max\{m_{\mathcal{T}}(S) + m_{\mathcal{T}}(T) - \underline{m_{\mathcal{T}}}(S \cap T); S, T, \subseteq N\ S \cup T = K$
card $S \setminus T =$ card $T \setminus S = 1\}$.
Then $m_{\mathcal{T}}$ is a convex multiset and $\mathcal{T} = \{A \in \mathcal{U}; m_{\mathcal{T}}(A) > 0\}$.

Proof By the definition of $m_{\mathcal{T}}$ it is easy to verify the condition (*ii*) from Lemma 2.7a. By Lemma 2.7b $m_{\mathcal{T}}$ is a multiset and $\mathcal{T} = \{A \in \mathcal{U}; m_{\mathcal{T}}(A) > 0\}$. ■

2.2.4 Base and Skeleton

The above mentioned necessary conditions will be used in Appendix A (see the third installment of the work, section 3.2) to show:

ASSERTION 2.2 The set of elementary imsets is the base of $\prec$.
Thus the base is known. What about the skeleton? It easily follows from Lemma 2.6b and Assertion 1.4(a), §1.3.4 that:

CONSEQUENCE 2.5 The skeleton for $\prec$ exists.
Thus, from the theoretical point of view we know how to recognize structural imsets. Namely: $0 \prec u$ iff $[\langle a, u\rangle \geq 0$ for each skeletal imset $a]$.
Nevertheless, the problem of practical finding of the structural skeleton is open. A convenient characterization of skeletal imsets giving an algorithm finding the skeleton for every number of attributes is needed. Note that in case card $N = 3$ the skeleton has 5 imsets while the base has 6 imsets (see Example 3.1, §3.1.1).
In [Studený, 1991] the skeleton in case card $N = 4$ is found. In this case, every skeletal imset has the form $m_{\mathcal{T}}$ for some ascending $\mathcal{T} \subseteq \mathcal{U}$ (see Lemma 2.8) and there exist exactly 37 skeletal structural imsets (while the base has 24 elements). The skeleton can be divided into 10 classes (every permutation of elements of N gives an "isomorphic" skeletal imset). The list of these classes is given (note that S, T, R, V are different subsets of a three-element set N) in Table 2.1.

2.3 STRUCTURAL FACES AS MODELS FOR CI-STRUCTURE

The aim of this section is to relate structural faces and imsets to dependency models. First, a dependency model is assigned to every structural face. This defines a one-

Table 2.1 Skeletal structural imsets in case card $N = 4$.

I	$m_{\mathcal{T}}$	where $\mathcal{T}_{min} = \{N\}$	1 representative
II	$m_{\mathcal{T}}$	where $\mathcal{T}_{min} = \{S\}$	4 representatives
III	$m_{\mathcal{T}}$	where $\mathcal{T}_{min} = \{S, T, R\}$	4 representatives
IV	$m_{\mathcal{T}}$	where $\mathcal{T}_{min} = \{S \cap T\}$	6 representatives
V	$m_{\mathcal{T}}$	where $\mathcal{T}_{min} = \{S, T, R, V\}$	1 representative
VI	$m_{\mathcal{T}}$	where $\mathcal{T}_{min} = \{S \cap T, R, V\}$	6 representatives
VII	$m_{\mathcal{T}}$	where $\mathcal{T}_{min} = \{S, T \cap R, T \cap V, R \cap V\}$	4 representatives
VIII	$m_{\mathcal{T}}$	where $\mathcal{T}_{min} = \{S \cap T, S \cap R, S \cap V\}$	4 representatives
IX	$m_{\mathcal{T}}$	where $\mathcal{T}_{min} = \{S \cap T, S \cap R, S \cap V, T \cap R, T \cap V\}$	6 representatives
X	$m_{\mathcal{T}}$	where $\mathcal{T}_{min} = \{S \cap T, S \cap R, S \cap V, T \cap R, T \cap V, R \cap V\}$	1 representative

to-one correspondence between structural faces and a certain class of semigraphoids also called structural. Then using the natural mapping of imsets into faces (see Lemma 2.5) a dependency model is assigned to every structural imset. Nevertheless, every probability measure induces a structural face through the concept of the multiinformation function and in this way it is shown that every probabilistically representable semigraphoid is structural. Some consequences concern facial implication; it is derived that semigraphoid derivability entails facial implication and facial implication entails probabilistic implication. Another consequence says that the description of CI-structures by means of structural imsets (or faces) and the description by means of dependency models are equivalent.

2.3.1 Primary Mapping

The first step to set up a connection between dependency models and structural faces is to define the primary mapping from the set of triplets $T_*(N)$ (see Def 1.3, §1.1.2) to semielementary imsets.

DEFINITION 2.10 (primary mapping)
Define a mapping i: $T_*(N) \to E_{sem}$ called the *primary mapping* as follows: to every triplet $\langle A, B, C\rangle \in T_*(N)$ assign $u \in E_{sem}$ such that

$$\bar{u} = \delta_{A\cup B\cup C} - \delta_{A\cup C} - \delta_{B\cup C} + \delta_C.$$

(Evidently, $\bar{u}$ satisfies (N.1)–(N.2)—see section 1.2.)

REMARK 2.2 (primary mapping is relatively injective)
Clearly, $i(\langle A, B, C\rangle) = i(\langle B, A, C\rangle)$ for every triplet $\langle A, B, C\rangle$. But this is the only reason why i is not injective. Using Lemma 1.2 (section 1.2) it is easy to see that

$$i(\langle A, B, C\rangle) = i(\langle A', B', C'\rangle) \text{ iff } \quad [C = C' \ \& \ \{A, B\} = \{A', B'\}].$$

Similarly, it is evident from the definitions that i maps $T_*(N)$ onto $E_{sem}\backslash\{0\}$.

Now, dependency models can be assigned to structural faces:

DEFINITION 2.11 (dependency model corresponding to a face)
For a face F with respect to the structural ordering the dependency model $i_{-1}(F)$ i.e. $\{t \in T_*(N);\ i(t) \in F\}$ will be called the *dependency model corresponding to F*.

LEMMA 2.9 The mapping $F \rightarrow i_{-1}(F)$ considered on the class of structural faces is injective. Moreover, every dependency model $i_{-1}(F)$ is a semigraphoid.

Proof I. Suppose that F_1 and F_2 are structural faces with $i_{-1}(F_1) = i_{-1}(F_2)$. Whenever $u \in E$ (elementary imset) then evidently $u \in E_{sem}\backslash\{0\}$ and by Remark 2.2 there exists $t \in T_*(N)$ with $u = i(t)$. Therefore:

$$u \in F_1 \Leftrightarrow t \in i_{-1}(F_1) \Leftrightarrow t \in i_{-1}(F_2) \Leftrightarrow u \in F_2.$$

Thus, $F_1 \cap E = F_2 \cap E$. As E is the base of the structural ordering (Assertion 2.2) by Theorem 2.1, §2.1.4 it distinguishes faces. Hence $F_1 = F_2$.

II. For a structural face F the symmetry condition for $i_{-1}(F)$ (see Def. 1.6, §1.1.3):
(a) $\langle A, B, C\rangle \in i_{-1}(F) \Leftrightarrow \langle B, A, C\rangle \in i_{-1}(F)$
follows directly from Remark 2.2. To prove the second condition
(b) $\langle A, BC, D\rangle \in i_{-1}(F) \Leftrightarrow \langle A, B, CD\rangle, \langle A, C, D\rangle \in i_{-1}(F)$
observe that $\langle A, C, D\rangle \in i_{-1}(F) \Leftrightarrow i(\langle A, C, D\rangle) \in F$ i.e. $\delta_{ACD} - \delta_{AD} - \delta_{CD} + \delta_D$ restricted to $\mathcal{U}$ belongs to F and similarly for the other triplets. But the equality $(\delta_{ABCD} - \delta_{AD} - \delta_{BCD} + \delta_D) = (\delta_{ABCD} - \delta_{ACD} - \delta_{BCD} + \delta_{CD}) + (\delta_{ACD} - \delta_{AD} - \delta_{CD} + \delta_D)$ restricted to $\mathcal{U}$ enables us to derive the implication $\Leftarrow$ in (b) by composition for F and the implication $\Rightarrow$ in (b) by decomposition (both summands belong to E_{sem}, and therefore are structural imsets by Theorem 2.4a, §2.2.3). ■

The following easy consequence relates semigraphoid derivability with facial implication.

CONSEQUENCE 2.6 Whenever $I \subseteq T_*(N)$ and $t \in T_*(N)$ then

$$I \vdash_{\text{sem}} t \quad \text{entails} \quad i(I) \mapsto i(t).$$

Proof Consider a face F with $i(I) \subseteq F$; it suffices to derive $i(t) \in F$ (see Def 2.3). The assumption says that there exists a derivation sequence $k_1, \ldots, k_n = t \subseteq T_*(N)$ (see Def 1.7, §1.1.3). It is straightforward to prove by induction that $\forall j = 1, \ldots, n \quad k_j \in i_{-1}(F)$ (by Lemma 2.9 $i_{-1}(F)$ is closed under all semigraphoid inference rules). ■

2.3.2 Structural Semigraphoids

Proceeding with the development of the theory, we introduce:

DEFINITION 2.12 (structural semigraphoid)
Every dependency model of the form $i_{-1}(F)$ where F is a structural face is called a *structural semigraphoid* (by Lemma 2.9 it is indeed a semigraphoid).

The adjective "structural" is indeed meaningful as there exist semigraphoids which are not structural. An example follows.

EXAMPLE 2.1 (nonstructural semigraphoid)
Consider $N = \{0, 1, 2, 3\}$ and put[10]:

$$I = \{\langle 0, 1, 2\rangle, \langle 1, 0, 2\rangle, \langle 0, 2, 3\rangle, \langle 2, 0, 3\rangle, \langle 0, 3, 1\rangle, \langle 3, 0, 1\rangle\}.$$

[10] $\langle a, b, c\rangle$ is written instead of $\langle\{a\}, \{b\}, \{c\}\rangle$.

Clearly, it is a semigraphoid. Suppose that $I \subseteq i_{-1}(F)$ for a structural face F. For our purpose, it suffices to show that $\langle 0, 2, 1\rangle \in i_{-1}(F)$. By Def 2.11 the following imsets belong to F:

$$i(\langle 0, 1, 2\rangle) = \delta_{\{0,1,2\}} - \delta_{\{0,2\}} - \delta_{\{1,2\}}$$

$$i(\langle 0, 2, 3\rangle) = \delta_{\{0,2,3\}} - \delta_{\{0,3\}} - \delta_{\{2,3\}}$$

$$i(\langle 0, 3, 1\rangle) = \delta_{\{0,1,3\}} - \delta_{\{0,1\}} - \delta_{\{1,3\}}.$$

Hence, by (F.1) their sum u belongs to F. To derive by (F.2) that $i(\langle 0, 2, 1\rangle) = \delta_{\{0,1,2\}} - \delta_{\{0,1\}} - \delta_{\{1,2\}} \in F$ write:

$$u - (\delta_{\{0,1,2\}} - \delta_{\{0,1\}} - \delta_{\{1,2\}}) = \{\delta_{\{0,2,3\}} - \delta_{\{0,2\}} - \delta_{\{2,3\}}\} + \{\delta_{\{0,1,3\}} - \delta_{\{0,3\}} - \delta_{\{1,3\}}\}.$$

The right-hand side is a sum of elementary imsets, and is therefore a structural imset.

REMARK 2.3 (struct. finite semigraphoids have no finite axiomatic characterization) Structural semigraphoids cannot be characterized as dependency models closed under a finite number of inference rules. Note that this can be proved in the same way as was used in [Studený, 1992] for the analogical result concerning probabilistically representable dependency models. In fact, every structural semigraphoid must be closed under all "inference rules" proved valid for probabilistically representable dependency models there (Proposition 1 in [Studený, 1992]) and since any probabilistically representable semigraphoid is structural (see below, Consequence 2.9, §2.3.4) the consideration from Consequence 1 in [Studený, 1992] can be repeated.

The above mentioned correspondence with dependency models can be transferred to structural imsets. Namely, every structural imset u determines the face $\bar{f}(\{u\})$ (see Def 2.2, §2.1.1). The corresponding dependency model $\{t \in T_*(N);\ i(t) \in \bar{f}(\{u\})\}$ can be expressed as follows (use Lemma 2.2(a)):

DEFINITION 2.13 (dependency model corresponding to imset)
For a structural imset $u \in Z(<)$ the dependency model $\{t \in T_*(N); u \mapsto i(t)\}$ i.e. $i_{-1}(\bar{f}\{u\})$ will be called the *dependency model corresponding to u* and denoted by I_u.

Note that I_u can be written in another form (use Lemma 2.2(b)):

FACT 2.5 $I_u = \{t \in T_*(N);\ \exists k \in \mathbb{Z}^+\ k \cdot u - i(t) \in Z(<)\}$ whenever $u \in Z(<)$.

By combining Lemma 2.9 and Lemma 2.5, §2.1.6 we get:

SUMMARY 2.3 The mapping $u \to I_u$ considered on structural imsets is relatively injective with respect to facial equivalence i.e.: $I_u = I_v$ iff $u \approx v$ whenever $u, v \in Z(<)$. Moreover, it maps structural imsets onto the class of structural semigraphoids.

2.3.3 Multiinformation Function

Faces and imsets will be related to probability measures by means of the concept of the multiinformation function having its source in information theory.

DEFINITION 2.14 (multiinformation function)
Given a probability measure P over N define its *multiinformation function* M : exp $N \to \mathbb{R}$ as follows:

$$M(\emptyset) = 0$$
$$M(S) = H(P^S; \prod_{i\in S} P^{\{i\}}) \quad \text{for} \quad \emptyset \neq S \subseteq V$$

where $H(R, Q)$ denotes the *relative entropy of* R with respect to Q defined by the formula

$$H(R, Q) = \sum_{R(x)>0} R(x) \cdot \ln (R(x)/Q(x))$$

where [$Q(x) = 0$ implies $R(x) = 0$]; this condition is satisfied for $R = P^S$, $Q = \Pi_{i\in S} P^{\{i\}}$.

Terminological remark Multiinformation generalizes the well-known information-theoretical concept of mutual information serving as a measure of dependence of two random variables. Multiinformation serves as a measure of stochastic dependence of two or more random variables. This view led me to adopt the name "multiinformation" in [Studený, 1989]. Another name, "entaxy", was used by Malvestuto [1983].

The following lemma summarizes the results from [Studený, 1989]§4,5 which allow the use of multiinformation as a tool for study of conditional independence.

LEMMA 2.10 For $P \in \mathcal{P}(N)$ (see Def 1.1, §1.1.1), its multiinformation function is a settled convex function (see Def 2.9). Moreover, whenever $\langle A, B, C\rangle \in T_*(N)$: P obeys $\langle A, B, C\rangle$ iff $M(ABC) - M(AC) - M(BC) + M(C) = 0$.

Usually, M will be considered as a function on $\mathcal{U}$ in the sequel. The principal concept follows.

DEFINITION 2.15 (probability measure complies with imset)
For $P \in \mathcal{P}(N)$ and $u \in Z(<)$ say that P *complies with* u iff $\langle M, u\rangle = 0$ where M is the multiinformation function of P (the scalar product $\langle \cdot, \cdot \rangle$ is defined in Def. 1.13b, §1.3.1).

Note that the second part of Lemma 2.10 can be reformulated as follows:

FACT 2.6 Whenever $P \in \mathcal{P}(N)$ and $t \in T_*(N)$ then P obeys t iff P complies with $i(t)$.

2.3.4 Induced Face

THEOREM 2.5 Suppose that P is a probability measure over N. Then the set $F_P = \{u \in Z(<); P \text{ complies with } u\}$ is a structural face.

Proof If M is the multiinformation function of P, then the set $F = \{u \in Z(<); \langle M, u\rangle = 0\}$ satisfies (F.0) and (F.1) owing to properties of scalar product. To show

(F.2) consider $u, v \in Z(<)$ with $u + v \in F$. Nevertheless, as E_{sem} establishes $<$ (Theorem 2.4a) $n \cdot u = \Sigma_{w \in E_{sem}} k_w \cdot w$ with $n \in \mathbb{N}$, $k_w \in \mathbb{Z}^+$. As M is convex (Lemma 2.10) $\langle M, n \cdot u \rangle = \Sigma_{w \in E_{sem}} k_w \langle M, w \rangle \geq 0$. Hence $\langle M, u \rangle \geq 0$ and similarly $\langle M, v \rangle \geq 0$. Thus $0 = \langle M, u + v \rangle = \langle M, u \rangle + \langle M, v \rangle$ gives $\langle M, u \rangle = \langle M, v \rangle = 0$ i.e. $u, v \in F$. ■

The previous theorem makes possible the definition:

DEFINITION 2.16 (face induced by probability measure)
For $P \in \mathcal{P}(N)$ the structural face $\{u \in Z(<); P$ complies with $u\}$ will be called the *face induced* by P and denoted by F_P.

Moreover, Theorem 2.5 has some important consequences. The first one relates the description of CI-structures by means of dependency models with complying with imsets. Therefore complying with a structural imset can be understood as certain type of description of the CI-structure.

CONSEQUENCE 2.7 For each $P \in \mathcal{P}(N)$ and $u \in Z(<)$ it holds that P complies with u iff I_u is a submodel of CI-structure of P.

Proof For $u \in F_P$ consider $t \in I_u$. To show that P obeys t observe that $i(t) \in F_P$ (since $u \mapsto i(t)$ and F_P is a face) and use Fact 2.6.
Conversely, suppose that P obeys every $t \in I_u$. Then by Theorem 2.4a write $n \cdot u = \Sigma_{w \in E_{sem}} k_w \cdot w$ where $n \in \mathbb{N}$, $k_w \in \mathbb{Z}^+$. For every $w \in E_{sem} \setminus \{0\}$ with $k_w \neq 0$ find $t_w \in T_*(N)$ with $w = i(t_w)$ (see Remark 2.2). By Fact 2.5 $t_w \in I_u$ and therefore P obeys t_w i.e. P complies with w (Fact 2.6). As P complies with every $w \in E_{sem} \setminus \{0\}$ with $k_w \neq 0$ and F_P is a face by (F.1) and (F.2), it follows that $u \in F_P$. ■

A further consequence relates facial implication with probabilistic implication.

CONSEQUENCE 2.8 Whenever $I \subseteq T_*(N)$ and $t \in T_*(N)$ then $i(I) \mapsto i(t)$ entails $I \models s$.

Proof Consider $P \in \mathcal{P}(N)$ which obeys I. By Fact 2.6 $\{i(k); k \in I\} \subseteq F_P$. As F_P is a face $i(I) \mapsto i(t)$ gives $i(t) \in F_P$ i.e. P obeys t by Fact 2.6. ■

Probabilistically representable dependency models can be characterized by means of faces as follows:

ASSERTION 2.3 For each $P \in \mathcal{P}(N)$ the dependency model corresponding to the face induced by P coincides with the model of CI-structure of P (see Def 1.4, §1.1.2).

Proof Write for $t \in T_*(N)$ using Fact 2.6, Def 2.16 and Def 2.11 : P obeys $t \Leftrightarrow P$ complies with $i(t) \in F_P \Leftrightarrow t \in i_{-1}(F_P)$. ■

This assertion has a significant corollary:

CONSEQUENCE 2.9 Every probabilistically representable dependency model is a structural semigraphoid.

Proof Combine Def 1.5 from §1.1.2, Assertion 2.3, Theorem 2.5 and Def 2.12. ■

2.3.5 Multiinformation Ordering

To make faces and imsets completely parallel with dependency models let us add the following definition (an analogue of Def 1.5 from §1.1.2).

DEFINITION 2.17 (probabilistically representable faces and imsets)
Suppose that $\Phi \subseteq \mathcal{P}(N)$. Say that a structural face F is *represented in* Φ iff $F = F_P$ for some $P \in \Phi$. A structural imset $u \in Z(\prec)$ is *represented in* Φ iff $\bar{f}(\{u\})$ is represented in Φ or equivalently I_u is represented in Φ (by Assertion 2.3). *A probabilistically representable* face (resp. imset) is a face (resp. imset) representable in $\mathcal{P}(N)$.

We conclude this section with the concept of multiinformation quasiordering discussed in the Conclusions (see the third part).

DEFINITION 2.18 (multiinformation quasiordering)
For $\Phi \subseteq \mathcal{P}(N)$ define the *multiinformation quasiordering corresponding to* Φ as the scalar product quasiordering (see Def 1.13c, §1.3.1) induced by the class of multiinformation functions of measures from Φ. By the *standard multiinformation ordering* is meant the multiinformation quasiordering corresponding to $\mathcal{P}(N)$.

In case that Φ is sufficiently large the corresponding quasiordering has to be an ordering. For example, the following condition is sufficient (this is the case of the standard multiinformation ordering):

$\forall A \in \mathcal{U}\ \exists P \in \Phi$ such that its multiinformation function has the form:

$$M(B) = \begin{cases} k > 0 & \text{in case } A \subseteq B \\ 0 & \text{otherwise.} \end{cases}$$

Certainly, it can be shown by Lemma 2.10 that the multiinformation ordering is weaker than the structural one. Nevertheless, so far I don't know whether they differ.

REFERENCES

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

Malvestuto, F. M. [1983], "Theory of random observables in relational data bases." *Information Systems*, **8**(4), pp. 281–289.

Rosenmüller, J. and H. G. Weidner [1974], "Extreme convex set functions with finite carrier: general theory." *Discrete Mathematics*, **10**(3/4), pp. 343–382.

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

Studený, M. [1991], "Convex set functions I and II." research reports n. 1733 and n. 1734, Institute of Information Theory and Automation, Prague, November 1991.

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

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