Transactions of the Eleventh Prague Conference on
Inform. Theory, Stat. Dec. Functions and Random Proc.
ACADEMIA: Prague 1992, Vol. B, 189–200

ASCENDING AND DESCENDING CONDITIONAL INDEPENDENCE RELATIONS

František Matúš

Prague

Keywords: Conditional independence, Connectedness, Matroids, Semimodular functions, Semigraphoids, Independence relations, Markov networks

ABSTRACT

We explore the conditional probabilistic independences of systems of random variables $(I, J|K)$, to read "I is conditionally independent of J given K", which are ascending $(I, J|K) \Rightarrow (I, J|K \cup L)$ or descending $(I, J|K \cup L) \Rightarrow (I, J|K)$. The resulting abstract independence structures can be equivalently described by weak families of connected sets. Using, in addition, probabilistic representations of matroids we present probabilistic representations of the structures and formulate duality conjectures. Local and global forms of semigraphoids and these independence structures are elaborated and special cases including unconditional independence relations, fixed–context relations and Markov networks are discussed.

1 INTRODUCTION

Let N be a fixed finite set, $\mathcal{P}(N)$ the family of its subsets, $\mathcal{R}(N)$ the family of all triples $(i, j|K)$, where $K \subset N$, $i, j \in N - K$, $i \neq j$, and $\mathcal{T}(N)$ the family of all triples $(I, J|K)$, where $I, J, K \subset N$ disjoint. We shall not distinguish between triples $(i, j|K)$ and $(j, i|K)$ ($(I, J|K)$ and $(J, I|K)$). To avoid clumsy expressions we shall often write singletons of N without braces. Subsets of $\mathcal{R}(N)$ ($\mathcal{T}(N)$) will be said to be local (global) relations on N.

By $\xi = (\xi_i)_{i \in N}$ we shall denote a system of random variables, every one taking only finite number of values, and by $\xi_K = (\xi_k)_{k \in K}$, $K \subset N$, the corresponding subsystem; $\xi_\emptyset$ is supposed to be constant. The fact that ξ_I is conditionally independent of ξ_J given ξ_K, $I, J, K \subset N$, will be in symbols written as $\xi_I \perp \xi_J \,|\, \xi_K$.

Definition 1. *A local relation $\mathcal{K}$ on N is called probabilistically (p-) representable if there exists ξ (a representation of $\mathcal{K}$) such that*

$$(i, j|K) \in \mathcal{K} \iff \xi_i \perp \xi_j \,|\, \xi_K.$$

The p–representability of global relations is determined analogically.

The problem of characterization of all p–representable global relations on N was formulated, in another language, by Pearl (1986) who presented these two elementary properties of every p–representable relation $\mathcal{M}$

$$\text{(CI 0)} \qquad J,\, K \subset N, \quad J \cap K = \emptyset \iff (\emptyset,\, J|K) \in \mathcal{M},$$
$$\text{(CI 1)} \qquad (I,\, J|K \cup L) \in \mathcal{M}, \quad (I,\, K|L) \in \mathcal{M} \iff (I,\, J \cup K|L) \in \mathcal{M}, \quad J \cap K = \emptyset$$

(see also Dawid (1979) and Mouchart, Rolin (1984)). The relations obeying these properties correspond to the semigraphoids, cf. Geiger, Pearl (1989). Studený (1989) found this additional property

$$\text{(CI 2)} \qquad \triangle(I, J;\, K, L|M) \subset \mathcal{M} \iff \triangle(K, L;\, I, J|M) \subset \mathcal{M},$$

where $\triangle(I, J;\, K, L|M) = \{(K, L|I \cup J \cup M),\ (K, L|M), (I, J|K \cup M),\ (I, J|L \cup M)\}$, and recently a countable class of "axioms" of this form (Studený (1990)). Hence, one would embark on a quest for another ways of characterization. Our work indicates that matroid theory, the classical theory of independence structures, would be a remedy.

Certain special subproblems have been settled up to now: The unconditional case (when one deals only with triples $(I,\, J|\emptyset)$) independently by Matúš (1988), Geiger, Paz, Pearl (1989) and the fixed–context one (i. e. the restriction to the triples $(I, J|K)$, $I + J + K = N$) by Geiger, Pearl (1989). In both cases the corresponding independence structures are essentially same and describe, moreover, other kinds of independences like algebraical one in lattices (Matúš (1990 b)), MVD and crosses in databases (Malvestuto (1990), Paredaens (1980)), etc.

The aim of the paper is to present two classes of p–representable relations namely of those which are ascending

$$\text{(A)} \qquad (I, J|K) \in \mathcal{M}, \qquad L \subset N - (I \cup J) \implies (I, J|K \cup L) \in \mathcal{M}$$

and descending

$$\text{(D)} \qquad (I, J|K \cup L) \in \mathcal{M} \implies (I, J|K) \in \mathcal{M}$$

(remark that the conditional independence structures of Markov networks, cf. Pearl (1986), represent a subclass of the first class). The structure of these classes is closely related to an abstract theory of connectedness (similarly like in Matúš (1990 b)). Further, we introduce the notion of p–representable matroids (section 3) and after proving that linear representability implies p–representability (Theorem 2) we can construct very easy representations of all relations holding (A), (CI 0), (CI 1), (CI 2). Duality of matroids leads not only to the representations of the descending relations but also to interesting duality conjectures (section 4).

All these results are worked out firstly for local relations (because of the brevity of explanation) and then translated into the equivalent global form (sections 5, 6). The local–global terminology is motivated by the similar one concerning the semimodular functions, cf. White (1986). At the end, unconditional, fixed–context and Markov versions of the problem are discussed as consequences of the results obtained (sections 7, 8, 9).

2 *ACI*–RELATIONS AND C^w–FAMILIES

Definition 2. *A subset $\mathcal{K}$ of $\mathcal{R}(N)$ is called (local) ACI–relation on N if it fulfils these three conditions*

(a) $$(i,j|K) \in \mathcal{K}, \quad L \subset N - (i \cup j) \implies (i,j|K \cup L) \in \mathcal{K},$$

(a1) $$(i,j|k \cup L) \in \mathcal{K}, \quad (i,k|L) \in \mathcal{K} \implies (i,j|L) \in \mathcal{K},$$

(a2) $$(i,j|k \cup L) \in \mathcal{K}, \quad (i,j|\ell \cup L) \in \mathcal{K}, \quad (k,\ell|L) \in \mathcal{K} \implies (i,j|L) \in \mathcal{K}.$$

We shall denote the class of all *ACI*–relations on N by $\mathbf{A}(N)$. It is plain that $\mathbf{A}(N)$ is a closure system on $\mathcal{R}(N)$; its closure operator α maps every $\mathcal{K} \subset \mathcal{R}(N)$ into the smallest *ACI*–relation $\alpha(\mathcal{K})$ on N containing $\mathcal{K}$. In this sense $\mathbf{A}(N)$ will be taken for a lattice (for all these notions see Birkhoff (1967), Ch. 5).

Definition 3. *A family $\mathcal{V}$ of subsets of N containing empty set and all signeltons of N is called C^w–family (<u>weak family of connected sets</u>) on N if it is closed under <u>double unions</u>; we have been proposing the last phrase in order to label the condition*

$$(\forall\ i,j \in M \ \ \exists\ L_{i,j} \in \mathcal{V}: \ \ i \cup j \subset L_{i,j} \subset M) \implies M \in \mathcal{V}.$$

This axiom of C^w–families issues from a well known lemma on connected sets in topological spaces "If to every two elements of a set M there exists a connected subset of M contaning them then M is connected alike", which dates back to Hausdorff (1914). The class $\mathbf{W}(N)$ of all C^w–families on N has, similarly like $\mathbf{A}(N)$, the structure of lattice; to its closure operator we reserve the symbol ω.

Theorem 1. *There is a one–to–one and onto correspondence between the classes $\mathbf{A}(N)$ and $\mathbf{W}(N)$ (these lattices are anti–isomorphical).*

Proof: Consider the binary relation $\circ \subset \mathcal{R}(N) \times \mathcal{P}(N)$ given by

$$(i,j|K) \ \circ \ M \iff M \cap (i \cup j \cup K) \neq i \cup j$$

and the Galois connection arising from it

$$\begin{aligned} \mathcal{K}^* &= \{M \in \mathcal{P}(N);\ \forall (i,j|K) \in \mathcal{K}:\ (i,j|K) \circ M\}, \qquad \mathcal{K} \subset \mathcal{R}(N), \\ \mathcal{W}^* &= \{(i,j|K) \in \mathcal{R}(N);\ \forall M \in \mathcal{W}:\ (i,j|K) \circ M\}, \qquad \mathcal{W} \subset \mathcal{P}(N). \end{aligned}$$

We shall prove that $\mathcal{K}^{**} = \alpha(\mathcal{K})$ for $\mathcal{K} \subset \mathcal{R}(N)$ and $\mathcal{W}^{**} = \omega(\mathcal{W})$ for $\mathcal{W} \subset \mathcal{P}(N)$. Having verified these assertions Theorem 1 follows from the general properties of Galois connections (cf. Birkhoff (1967) for more details), especially mappings $\mathcal{K} \to \mathcal{K}^*$ and $\mathcal{W} \to \mathcal{W}^*$ (restricted properly) are mutually inverse anti–isomorphisms of $\mathbf{A}(N)$ and $\mathbf{W}(N)$.

First, we write out some apparent properties of the relation $\circ$. Evidently, $(i,j|K) \circ \emptyset$ and $(i,j|K) \circ \ell$ if $(i,j|K) \in \mathcal{R}(N)$, $\ell \in N$. Further, if $M = \bigcup \mathcal{W}$ is the double union of $\mathcal{W} \subset \mathcal{P}(N)$ and $(i,j|K) \circ L$ for all $L \in \mathcal{W}$ then $(i,j|K) \circ M$. These facts imply that $\mathcal{K}^*$, $\mathcal{K} \subset \mathcal{R}(N)$, is always a C^w–family and $\mathcal{W}^* = (\omega(\mathcal{W}))^*$, $\mathcal{W} \subset \mathcal{P}(N)$. Similarly, from

$$(i,j|K) \circ M, \quad L \subset N - (i \cup j) \implies (i,j|K \cup L) \circ M,$$

$$(i,j|k\cup L)\circ M,\quad (i,k|L)\circ M \implies (i,j|L)\circ M,$$

$$(i,j|k\cup L)\circ M,\quad (i,j|\ell\cup L)\circ M,\quad (k,\ell|L)\circ M \implies (i,j|L)\circ M$$

we deduce $\mathcal{W}^* \in \mathbf{A}(N)$, $\mathcal{W} \subset \mathcal{P}(N)$, and $\mathcal{K}^* = (\alpha(\mathcal{K}))^*$, $\mathcal{K} \subset \mathcal{R}(N)$. Hence $\mathcal{K} \subset \alpha(\mathcal{K}) \subset \mathcal{K}^{**}$, $\mathcal{W} \subset \omega(\mathcal{W}) \subset \mathcal{W}^{**}$ and it rests to verify $\mathcal{K} = \mathcal{K}^{**}$, $\mathcal{K} \in \mathbf{A}(N)$, and $\mathcal{W} = \mathcal{W}^{**}$, $\mathcal{W} \in \mathbf{W}(N)$.

On the contrary, let $(i,j|K) \in \mathcal{K}^{**} - \mathcal{K}$ for an *ACI*–relation $\mathcal{K}$ on N and let $(i,j|K)$ was chosen so that $|K| = \max\{|L|;\ (k,\ell|L) \in \mathcal{K}^{**} - \mathcal{K}\}$. As $(i,j|K) \not\circ (N-K)$ we get $N-K \notin \mathcal{K}^*$, i.e. there exists a triple $(k,\ell|L) \in \mathcal{K}$ such that $(k,\ell|L) \not\circ (N-K)$. We see $k,\ \ell \notin K$ and even by means of (a) $L = K$ can be assumed. The equality $i \cup j = k \cup \ell$ bring in immediately the contradiction $(i,j|K) \in \mathcal{K}$. If $(i\cup j)\cap(k\cup\ell)$ is a singleton, e.g. $i = \ell$, then $(i,k|K) \in \mathcal{K}$ and $(i,j|k\cup K) \in \mathcal{K}$ by the choice of $(i,j|K)$, i.e. $(i,j|K) \in \mathcal{K}$ (use (a1)) contradicts the starting assumption. And finally, if $i,\ j,\ k,\ \ell$ are four different elements of $N-K$ we arrive from $(k,\ell|K) \in \mathcal{K}$, $(i,j|k\cup K) \in \mathcal{K}$ and $(i,j|\ell\cup K) \in \mathcal{K}$ (by (a2)) again at the same contradiction.

The case of C^w–families can be treated analogically, let $M \in \mathcal{W}^{**} - \mathcal{W}$, $\mathcal{W} \in \mathbf{W}(N)$ (necessarily $|M| \geq 2$). To every pair $i,j \in M$ we have $(i,j|N-M) \not\circ M$, then $(i,j|N-M) \notin \mathcal{W}^*$ follows and therefore there is $L_{i,j} \in \mathcal{W}$ containing $i\cup j$ and contained in M. Since the double union of the family $\{L_{i,j};\ i,j \in M\} \subset \mathcal{W}$ belongs to $\mathcal{W}$ and equals M we get the contradiction $M \in \mathcal{W}$. □

3 P–REPRESENTATIONS

Information theoretical concepts like entropy and multiinformation are known to be powerful tools for distinguishing the conditional independences among random variables.

Remind that Shannon entropy of a random variable taking finite number of values with positive probabilities p_s $(\sum p_s = 1)$ is given by the expression

$$-\sum p_s \log p_s$$

If $\xi = (\xi_i)_{i\in N}$ is a system of random variables we shall write $H_\xi(I)$, $I \subset N$, for the entropy of the subsystem ξ_I. H_ξ is a real valued function on $\mathcal{P}(N)$ which is normalized $(H_\xi(\emptyset) = 0)$, nondecreasing $(I \subset J \subset N \ \Rightarrow\ H_\xi(I) \leq H_\xi(J))$ and globally semimodular

$$H_\xi(I) + H_\xi(J) \geq H_\xi(I\cap J) + H_\xi(I\cup J), \qquad I,\ J \subset N$$

(for proofs see Martin, England (1981)). The well known assertion, called here entropy balance rule, can be now stated as

$$\xi_I \perp \xi_J \,|\, \xi_K \iff H_\xi(I\cup K) + H_\xi(J\cup K) \ = \ H_\xi(K) + H_\xi(I\cup J\cup K), \qquad I,\ J,\ K \subset N.$$

One observes easily from it that intersection of two p–representable relations is also p–representable. In fact, if $\mathcal{K} \subset \mathcal{R}(N)$ is representable by ξ and $\mathcal{L} \subset \mathcal{R}(N)$ by η then we can handle ξ and η as two stochastically independent systems and create the system of random pairs $(\xi,\eta) = ((\xi_i,\eta_i))_{i\in N}$. From $H_{(\xi,\eta)} = H_\xi + H_\eta$ we deduce that $\mathcal{K}\cap\mathcal{L}$ is representable by (ξ,η).

Although this is a very challenging subject, we do not intend to explore the function H_ξ here. Only the definition and the theorem below are introduced. For the definition of matroids and further elementary facts we refer the reader to Welsh (1976) or White (1986).

Definition 4. *A matroid (N, r) on the set N with the rank function r is called p–representable if there exists ξ such that H_ξ equals r up to a positive multiplicative constant.*

Theorem 2. *If a matroid (N, r) is linearly representable over a finite field then it is p–representable.*

Proof: Recall that r is an integer–valued function on N which is nondecreasing, semimodular and $0 \le r(I) \le |I|$, $I \subset N$. The linear representation of a matroid (N, r) is a system $x = (x_i)_{i \in N}$ of vectors of a linear space $E^\rceil$ such that $r(I)$ equals the dimension of the linear span of the subfamily $x_I = (x_i)_{i \in I}$, $I \subset N$. Assume, in addition, that the dimension of $E^\rceil$ is $r(N)$. Now we interpret the space $E^\rceil$ as the dual space of an $r(N)$–dimensional linear space E so that x_i, $i \in N$, are linear functionals on E.

In the case of a finite ground field F we can consider E to be the probability space with the uniform probability $|F|^{-r(N)}$ of all its vectors – atoms and $\xi = x$ (i.e. $\xi_i = x_i$, $i \in N$) to be a system of random variables on E every one taking its values in F. For $I \subset N$, the subsystem ξ_I induces on E a partition blocks of which are $r(N) - r(I)$ dimensional parallel flats. Hence the atoms of ξ_I have the same probability $|F|^{-r(I)}$. The entropy of ξ_I is thus $r(I)$ having chosen $|F|$ for the base of the logarithm. □

The proof of the following assertion is now considerably short.

Theorem 3. *Every ACI–relation on N is p–representable.*

Proof: By the basic properties of Galois connections we have

$$\mathcal{K} = \mathcal{K}^{**} = \left(\bigcup \{\{L\};\ L \in \mathcal{K}^*\}\right)^* = \bigcap \{\{L\}^*;\ L \in \mathcal{K}^*\}$$

and consequently it suffices to prove that every *ACI*–relation $\{L\}^*$, $L \subset N$ is p–representable. The matroid (N, r_L) where

$$r_L(I) = \min\{1, |I \cap L|\}, \qquad I \subset N,$$

is easily recognized to be linearly representable over $GF(2)$. Hence, by Theorem 2, it is p–representable. The observation

$$\{L\}^* = \{(i, j|K) \in \mathcal{R}(N);\ r_L(i \cup K) + r_L(j \cup K) = r_L(K) + r_L(i \cup j \cup K)\}$$

together with the entropy balance rule close the proof. □

Taking regard on (CI 1) and (CI 2) we assure that any (local) p–representable relation $\mathcal{K}$ on N satisfies

(ci 1) $$(i, j|k \cup L) \in \mathcal{K},\ (i, k|L) \in \mathcal{K} \iff (i, k|j \cup L) \in \mathcal{K}, \quad (i, j|L) \in \mathcal{K},$$

(ci 2) $$\triangle(i, j;\, k, \ell|M) \subset \mathcal{K} \iff \triangle(k, \ell;\, i, j|M) \subset \mathcal{K}.$$

If $\mathcal{K}$ is moreover ascending then it obeys also (a1) and (a2), i.e. $\mathcal{K} \in \mathbf{A}(N)$. This gives us a justification for the interpretation of the abbreviation *ACI* as ascending conditional independence.

4 DUALITY OF MATROIDS AND *DCI*–RELATIONS

Theorem 2 enables to construct a large class of p–representable relations. Let (M, r) be a linearly representable (over a finite field) matroid and $\sum_{i \in N} M_i = M$ be a partition of M. Then

$$\mathcal{K} = \{(i,j|K) \in \mathcal{R}(N);\ r(M_{i\cup K}) + r(M_{j\cup K}) = r(M_K) + r(M_{i\cup j\cup K})\},$$

where $M_I = \sum_{i\in I} M_i$, $I \subset N$, is, owing to the entropy balance rule and Theorem 2, a p–representable relation on N.

The notion of the dual matroid $(M, r^{\rceil})$ of (M, r) with the rank function

$$r^{\rceil}(L) = |L| + r(M-L) - r(M), \qquad L \subset M,$$

brings in an additional symmetry. Since the dual matroid is linearly representable over the same field,

$$\mathcal{L} = \left\{(i,j|K) \in \mathcal{R}(N);\ r^{\rceil}(M_{i\cup K}) + r^{\rceil}(M_{j\cup K}) = r^{\rceil}(M_K) + r^{\rceil}(M_{i\cup j\cup K})\right\}$$

is also a p–representable relation. The notations $(i,j|K)^{\rceil} = (i,j|N-(i\cup j\cup K))$ and $\mathcal{K}^{\rceil} = \{(i,j|K)^{\rceil};\ (i,j|K) \in \mathcal{K}\}$ yield $\mathcal{K}^{\rceil} = \mathcal{L}$.

There are two natural questions to put forth. *Is the dual of a p–representable matroid always p–representable? Is the dual $\mathcal{K}^{\rceil}$ of a p–representable relation $\mathcal{K}$ on N p–representable too?*

Nevertheless, we can apply successfully these considerations to *ACI*–relations.

Definition 5. *A subset $\mathcal{K}$ of $\mathcal{R}(N)$ is called DCI–relation on N if it satisfies*

(d) $\qquad (i,j|K\cup L) \in \mathcal{K} \implies (i,j|K) \in \mathcal{K},$
(d1) $\qquad (i,j|k\cup L) \in \mathcal{K}\ (i,k|L) \in \mathcal{K} \implies (i,k|j\cup L) \in \mathcal{K},$
(d2) $\qquad (i,j|k\cup L) \in \mathcal{K}\ (i,j|\ell\cup L) \in \mathcal{K},\ (k,\ell|i\cup j\cup L) \in \mathcal{K} \implies (i,j|k\cup \ell\cup L) \in \mathcal{K}.$

It is easy to verify that $\mathcal{K}$ obeys (d), (d1) and (d2) if and only if $\mathcal{K}^{\rceil}$ does (a), (a1) and (a2), respectively, and vice versa. These facts can be commented e.g. as (a) is dual of (d). The mapping $\mathcal{K} \to \mathcal{K}^{\rceil}$ is thus a self-inverse, one–to–one and onto correspondence (isomorphism) between the class (lattice) $\mathbf{A}(N)$ of all *ACI*–relations and the class $\mathbf{D}(N)$ of all *DCI*–relations on N. Remark that due to Theorem 1 there is a bijection between $\mathbf{D}(N)$ and $\mathbf{W}(N)$ (explicitly set by a Galois connection arising from the binary relation

$$(i,j|K)\Box M \iff i \notin M \text{ or } j \notin M \text{ or } M-(i\cup j\cup K) \neq \emptyset).$$

Theorem 4. *Every DCI–relation on N is p–representable.*

Proof: We proceed like in the foregoing proof. For $\mathcal{K} \in \mathbf{D}(N)$ one have

$$\mathcal{K} = \left(\mathcal{K}^{\rceil *}\right)^{*\rceil} = \left(\bigcap\left\{\{L\}^*;\ L \in \mathcal{K}^{\rceil *}\right\}\right)^{\rceil} = \bigcap\left\{\{L\}^{*\rceil};\ L \in \mathcal{K}^{\rceil *}\right\}$$

and having commented the p–representability of *DCI*–relations $\{L^*\}^{\rceil}$ we see that $\mathcal{K}$ has a p–representation too. □

A comparision between (ci 1), (ci 2) and (d), (d1), (d2) yields that every p–representable descending relation on N belongs to $\mathbf{D}(N)$ and from now on we may read *DCI* as descending conditional independence.

5 SEMIGRAPHOIDS

The second part of the paper contains applications of the presented results. We feel that it is overloaded by various technicalities but we cannot evade them.

For $\mathcal{K} \subset \mathcal{R}(N)$ we set

$$\begin{aligned} \mathrm{gl}\,(\mathcal{K}) = \{(I,J|K) \in \mathcal{T}(N);\ \forall\ i \in I\ \forall\ j \in J\ \forall M \subset N : \\ K \subset M \subset (I-i) \cup (J-j) \cup K \implies (i,j|M) \in \mathcal{K}\} \end{aligned}$$

and for $\mathcal{M} \subset \mathcal{T}(N)$ simply $\mathrm{loc}\,(\mathcal{M}) = \mathcal{M} \cap \mathcal{R}(N)$.

Proposition 1. *The mappings* gl *and* loc *form a pair of mutually inverse bijections between the class of all local relations obeying (ci 1) and the class of all global relations obeying (CI 0) and (CI 1).*

Proof: For any $\mathcal{K} \subset \mathcal{R}(N)$ $\mathrm{gl}\,(\mathcal{K})$ fulfils evidently (CI 0), the implication $\Leftarrow$ in (CI 1) and $\mathcal{K} = \mathrm{loc}\,(\mathrm{gl}(\mathcal{K}))$. We shall first pay heed to the reverse implication.

Let $\mathcal{K}$ obeys (ci 1) and $(I,j|K \cup L) \in \mathrm{gl}\,(\mathcal{K})$, $(I,K|L) \in \mathrm{gl}\,(\mathcal{K})$. If $(i,j|M) \notin \mathcal{K}$ for $i \in I$ and $L \subset M \subset (I-i) \cup K \cup L$ then for all $k \in K - M$ also $(i,j|k \cup M) \notin \mathcal{K}$ (otherwise knowing that $(i,k|M) \in \mathcal{K}$ we get by (ci 1) $(i,j|M) \in \mathcal{K}$) and consequently $(i,j|K \cup M) \notin \mathcal{K}$ what is impossible. Hence $(i,j|M) \in \mathcal{K}$ as soon as $i \in I$ and $L \subset M \subset (I-i) \cup K \cup L$. If $i \in I$, $k \in K$, $L \subset M \subset (I-i) \cup (K-k) \cup L$ then trivially $(i,k|M) \in \mathcal{K}$. If $i \in I$, $k \in K$, $j \cup L \subset M \subset (I-i) \cup (K-k) \cup j \cup L$ then from $(i,k|M-j) \in \mathcal{K}$ and $(i,j|k \cup (M-j)) \in \mathcal{K}$ we conclude again by (ci 1) that $(i,k|M) \in \mathcal{K}$. We have just proved $(I, j \cup K|L) \in \mathrm{gl}\,(\mathcal{K})$, i.e. the implication $\Rightarrow$ in (CI 1) for arbitrary possible I, K, L and $J = j$.

Further, we can proceed by induction on the cardinality of J. The suitable idea is to partition $J = J_1 + J_2$ into two nonempty blocks and to employ the implication $\Leftarrow$ in (CI 1) and the induction assumption. Thus, from $(I,J|K \cup L) \in \mathrm{gl}\,(\mathcal{K})$ we get $(I,J_1|K \cup L) \in \mathrm{gl}\,(\mathcal{K})$ and $(I,J_2|J_1 \cup K \cup L) \in \mathrm{gl}\,(\mathcal{K})$ and moreover from $(I,K|L) \in \mathrm{gl}\,(\mathcal{K})$ evidently $(I, J_1 \cup K|L) \in \mathrm{gl}\,(\mathcal{K})$ and $(I, J_1 + J_2|K \cup L) \in \mathrm{gl}\,(\mathcal{K})$.

On the other hand, let $\mathcal{M} \subset \mathcal{T}(N)$ satisfies (CI 0) and (CI 1). Then $\mathrm{loc}\,(\mathcal{M})$ obeys plainly (ci 1) and it rests to verify $\mathrm{gl}\,(\mathrm{loc}(\mathcal{M})) = \mathcal{M}$, i.e.

$$(I,J|K) \in \mathcal{M} \iff [\forall\ i \in I\ \forall\ j \in J\ \forall\ M \subset N :\ K \subset M \subset (I-i) \cup (J-j) \cup K \implies (i,j|M) \in \mathcal{M}].$$

The nontrivial implication $\Leftarrow$ can be treated by induction on $|I+J|$. The case $|I| \leq 1$ and $|J| \leq 1$ is trivial. Loosing no generality we can assume $|I| \geq 1$, $J = J_1 + J_2$, J_1, J_2 nonempty. The induction argument leads to $(I,J_1|K) \in \mathcal{M}$ and to $(I,J_2|J_1 \cup K) \in \mathcal{M}$, and finally, (CI 1) to $(I,J|K) \in \mathcal{M}$. □

Proposition 1 brings in an equivalent local description of semigraphoids. It offers to view a semigraphoid $\mathcal{M}$ as a simple graph with vertices from N and edges evaluated by families of subsets of N, i.e. the edge ij is evaluated by the family $\{K \subset N;\ (i,j|K) \in \mathcal{M}\}$.

Going through the proof of Proposition 1 one concludes that the p–representability problems of local and global relations are equivalent.

6 GLOBAL *ACI*– AND *DCI*–RELATIONS

Definition 6. *A subset $\mathcal{M}$ of $\mathcal{T}(N)$ is called global ACI–relation (DCI–relation) on N if it fulfils (CI 0), (CI 1), (CI 2) and (A) ((D)).*

We need additional properties of local *ACI*–relations.

Lemma 1. *Let $\mathcal{K}$ be a local ACI–relation on N. Then*

$$[(i,j|K \cup L) \in \mathcal{K},\ \forall\ k \in K:\ (i,k|L) \in \mathcal{K}] \implies (i,j|L) \in \mathcal{K},$$

$$[(i,j|K \cup M) \in \mathcal{K},\ (i,j|L \cup M) \in \mathcal{K},\ \forall\ k \in K\ \forall \ell \in L:\ (k,\ell|M) \in \mathcal{K}] \implies (i,j|M) \in \mathcal{K}.$$

Proof: The first implication can be verified easily by induction on $|K|$. We approach immediately the second one. Employing induction on $|K+L|$, $K \neq \emptyset$, $L \neq \emptyset$ we see that is coincides with (a2) if $|K+L| = 2$.

Let $\ell \in L$, $K_0 \subset K$ and $(i,j|K_0 \cup (L-\ell) \cup M) \in \mathcal{K}$ (e.g. for $K_0 = K$ it is valid by (a)). Then for any $k \in K_0$ the three triples $(i,j|k \cup (K_0 - k) \cup (L-\ell) \cup M)$, $(i,j|\ell \cup (K_0-k) \cup (L-\ell) \cup M)$ and $(k,\ell|(K_0-k) \cup (L-\ell) \cup M)$ belong to $\mathcal{K}$ and due to (a2) also $(i,j|(K_0-k) \cup (L-\ell) \cup M) \in \mathcal{K}$. Consequently $(i,j|(L-\ell) \cup M) \in \mathcal{K}$ and the induction assumption gives $(i,j|M) \in \mathcal{K}$. □

Mappings gl and loc are mutually inverse bijections between the classes of local and global *ACI*–relations on N. In fact, if $\mathcal{M}$ is a global *ACI*–relation then clearly $\mathrm{loc}(\mathcal{M}) \in \mathbf{A}(N)$ (cf. Proposition 1) and if $\mathcal{K}$ is a local *ACI*–relation then

$$\mathrm{gl}(\mathcal{K}) = \{(I,J|K) \in \mathcal{T}(N);\ \forall\ i \in I\ \forall\ j \in J:\ (i,j|K) \in \mathcal{K}\}$$

satisfies (CI 0), (CI 1) (see Proposition 1), trivially (A) and by Lemma 1 also (CI 2).

Let us denote $\mathcal{M}^{\rceil} = \{(I,J|K) \in \mathcal{T}(N);\ (I,J|N-(I \cup J \cup K)) \in \mathcal{M}\}$ for any $\mathcal{M} \in \mathcal{T}(N)$. It is easy to find out that every one from the properties (CI 0), (CI 1) and (CI 2) is valid for $\mathcal{M}$ if it is valid for $\mathcal{M}^{\rceil}$; we should say they are selfdual. Also, (A) is dual of (D). We deduce now that the duality mapping $\mathcal{M} \to \mathcal{M}^{\rceil}$ is a selfinverse bijection between the classes of global *ACI*– and global *DCI*–relations. Let us remark that $(\mathrm{gl}(\mathcal{K}))^{\rceil} = \mathrm{gl}(\mathcal{K}^{\rceil})$, $\mathcal{K} \subset \mathcal{R}(N)$, takes place and this entails that gl and loc are bijections between the classes of local and global *DCI*–relations.

We can now close the problem of p–representations of ascending and descending relations by these statements.

Theorem 5. *The class of all local/global ACI–/DCI–relations on N is identical with the class of all local/global ascending/descending p–representable relations on N.*

7 *I*–RELATIONS

Let $\mathcal{Q}(N)$ be the family of all unordered pairs (I,J) of disjoint subsets of N. We shall discuss here an embedding of the class of *I*–relations (defined below) into the class of *DCI*–relations.

Definition 7. *A subset $\mathcal{L}$ of $\mathcal{Q}(N)$ is called I–relation on N if (cf. Matúš (1988))*

$$I \subset N \implies (\emptyset, I) \in \mathcal{L}, \tag{I0}$$

$$(I \cup J, K) \in \mathcal{L} \implies (I, K) \in \mathcal{L}, \tag{I1}$$

$$(I \cup J, K) \in \mathcal{L},\ (I, J) \in \mathcal{L} \implies (I \cup K, J) \in \mathcal{L}. \tag{I2}$$

For $\mathcal{L} \subset \mathcal{Q}(N)$ we set $\mathrm{con}(\mathcal{L}) = \{(i,j|I \cup J) \in \mathcal{R}(N);\ (i \cup I,\ j \cup J) \in \mathcal{L}\}$ and, reversely, for $\mathcal{K} \subset \mathcal{R}(N)$ $\mathrm{unc}(\mathcal{K}) = \{(I,J) \in \mathcal{Q}(N);\ \forall\ i \in I\ \forall\ j \in J:\ (i,j|(I-i) \cup (J-j)) \in \mathcal{K}\}$.

Proposition 2. *The mappings* con *and* unc *are mutually inverse bijections between the class of all $\mathcal{L} \subset \mathcal{Q}(N)$ satisfying (I0) and*

$$(I \cup J,\ K \cup L) \in \mathcal{L},\ (I \cup K,\ J \cup L) \in \mathcal{L} \implies (I, J \cup K \cup L) \in \mathcal{L} \tag{B}$$

and the class of all $\mathcal{K} \subset \mathcal{R}(N)$ satisfying

$$(i,j|k \cup L) \notin \mathcal{K},\ (j,k|i \cup L) \notin \mathcal{K} \implies (i,k|j \cup L) \notin \mathcal{K}. \tag{b}$$

Proof: We follow the same lines like in the proof of Proposition 1. If $\mathcal{L} \subset \mathcal{Q}(N)$ and $(i,k|j \cup L) \in \mathrm{con}(\mathcal{L})$ then there is a partition $I + K$ of L such that $(i \cup I,\ j \cup k \cup K) \in \mathcal{L}$ or $(i \cup j \cup I,\ k \cup K) \in \mathcal{L}$ and thus $(i,j|k \cup L) \in \mathrm{con}(\mathcal{L})$ or $(j,k|i \cup L) \in \mathrm{con}(\mathcal{L})$, i.e. (b) holds for $\mathrm{con}(\mathcal{L})$. It is also easy to see that $\mathrm{unc}(\mathcal{K})$ fulfils (I0) and (B) for any $\mathcal{K} \subset \mathcal{R}(N)$.

Now we are going to demonstrate the equality $\mathcal{K} = \mathrm{con}(\mathrm{unc}(\mathcal{K}))$ for $\mathcal{K}$ holding (b). Knowing that

$$\begin{aligned}&\mathrm{con}(\mathrm{unc}(\mathcal{K})) = \{(i,j|M) \in \mathcal{R}(N);\\ &\exists\ I + J = M\ \forall\ k \in i \cup I\ \forall\ \ell \in j \cup J:\ (k,\ell|(i \cup j \cup I \cup J) - (k \cup \ell)) \in \mathcal{K}\}\end{aligned}$$

the inclusion $\supset$ is trivial. On the other hand, let $(i,j|M) \in \mathcal{K}$ and $I = \{k \in M; (k,j|(i \cup M) - k) \in \mathcal{K}\}$, $J = M - I$. Then for all $k \in i \cup I$ and $\ell = j$ $(k,\ell|(i \cup M) - k)$ belongs to $\mathcal{K}$. In addition, for all $k \in i \cup I$ and $\ell \in J$ $(\ell,j|(i \cup M) - \ell) \notin \mathcal{K}$ and owing to (b) $(k,\ell|(i \cup j \cup M) - (k \cup \ell)) \in \mathcal{K}$ so that the inclusion $\subset$ takes place.

Arbitrary $\mathcal{L} \subset \mathcal{Q}(N)$ is contained in

$$\begin{aligned}&\mathrm{unc}(\mathrm{con}(\mathcal{L})) = \{(I,J) \in \mathcal{Q}(N);\\ &\forall\ i \in I,\ \forall\ j \in J\ \exists\ K + L = (I-i) \cup (J-j):\ (i \cup K,\ j \cup L) \in \mathcal{L}\}.\end{aligned}$$

We shall prove that $\mathcal{L}$ holding (I0) and (B) equals $\mathrm{unc}(\mathrm{con}(\mathcal{L}))$. If $(I,J) \in \mathrm{unc}(\mathrm{con}(\mathcal{L}))$ then for every $i \in I$ and $j \in J$ there is a pair $(K_{i,j}, L_{i,j}) \in \mathcal{L}$ such that $i \in K_{i,j}$, $j \in L_{i,j}$ and $I + J = K_{i,j} + L_{i,j}$. For all $i \in I$ let $M_i = \bigcap\{K_{i,j};\ j \in J\}$. Employing (B) we find $(M_i,\ (I \cup J) - M_i) \in \mathcal{L},\ i \in I$, however $i \in M_i \subset I$ and hence $\bigcup_{i \in I} M_i = I$. Proceeding again by (B) we arrive at $(I,J) \in \mathcal{L}$. □

Proposition 3. *The mappings* con *and* unc *establish mutually inverse bijections between the class $\mathbf{I}(N)$ of all I–relations on N and the subclass of $\mathbf{D}(N)$ determined by (b).*

Proof: If $\mathcal{K}$ is a local *DCI*–relation holding (b) then $\mathrm{unc}(\mathcal{K}) = \{(I,J) \in \mathcal{Q}(N);\ (I,J|\emptyset) \in \mathrm{gl}(\mathcal{K})\}$ and according to Proposition 1 the properties (CI 0) and (CI 1) of $\mathrm{gl}(\mathcal{K})$ imply the properties (I0), (I1) and (I2) of $\mathrm{unc}(\mathcal{K})$, i.e. $\mathrm{unc}(\mathcal{K}) \in \mathbf{I}(N)$.

If $\mathcal{L} \in \mathbf{I}(N)$ we shall demonstrate $\text{con}(\mathcal{L}) \in \mathbf{D}(N)$. The property (d) of $\text{con}(\mathcal{L})$ is implied by (I1). We proceed by the verification of (d1) for $\text{con}(\mathcal{L})$. Having $(i,j|k \cup L) \in \text{con}(\mathcal{L})$ then for some partition $M + J$ of L we have $(i \cup M,\ j \cup k \cup J) \in \mathcal{L}$ (but then immediately $(i,k|j \cup L) \in \text{con}(\mathcal{L})$) or, as we shall assume, $(i \cup k \cup M,\ j \cup J) \in \mathcal{L}$. If, moreover, $(i,k|L) \in \text{con}(\mathcal{L})$, i.e. $(i \cup I,\ k \cup K) \in \mathcal{L}$, where $I + K = L$, then by (I1) and (I2) $(i \cup (I \cap M), (j \cup k \cup L) - (I \cap M)) \in \mathcal{L}$ and hence again $(i,k|j \cup L) \in \text{con}(\mathcal{L})$. At the end we mention that (d1) and (b) imply always (d2). Since $\text{con}(\mathcal{L})$ satisfies (b) (by Proposition 2 and by the observation that I–relations obey (B)) it satisfies also (d2), i.e. $\text{con}(\mathcal{L}) \in \mathbf{D}(N)$. □

8 C–FAMILIES

In the preceding section we found the embedding con of the class $\mathbf{I}(N)$ into the class $\mathbf{D}(N)$. The duality mapping enables alternatively to embed $\mathbf{I}(N)$ into $\mathbf{A}(N)$ by the mapping $\mathcal{L} \to (\text{con}(\mathcal{L}))^{\rceil}$. The corresponding subclass of $\mathbf{A}(N)$ is determined by the dual property of (b)

$$\text{(t)} \qquad (i,k|L) \notin \mathcal{K},\ (j,k|L) \notin \mathcal{K} \implies (i,j|L) \notin \mathcal{K}.$$

This subclass can be characterized equivalently (beside the properties (a), (a1), (t)) by means of connected sets as we show here.

Definition 8. *A family $\mathcal{V}$ of subsets of N containing empty set and all singletons of N is called C–family (family of connected sets) on N if it contains together with any $I, J \subset N$ nondisjoint also their union $I \cup J$.*

Resemblance with the notions of connection in topology and in matroid theory is more than transparent: "The union of two connected sets with nonempty intersection is connected" (Hausdorff (1914)). Evidently the class $\mathbf{V}(N)$ of all C–families on N is a proper subclass of the class $\mathbf{W}(N)$ of all C^w–families on N as soon as $|N| \geq 3$.

Proposition 4. *A local ACI–relation $\mathcal{K}$ on N obeys the condition (t) if and only if $\mathcal{K}^*$ is a C–family.*

Proof: Let us suppose $I, J \in \mathcal{K}^*$ for $\mathcal{K} \in \mathbf{A}(N)$ holding (t) and $k \in I \cap J$. For every $i \in I - J$ and $j \in J - I$ $(i,k|N - (I \cup J)) \,ø\, I$ and $(j,k|N - (I \cup J) \,ø\, J$ take place and then by means of Theorem 1 $(i,k|N - (I \cup J)) \notin \mathcal{K} = \mathcal{K}^{**}$ and $(j,k|N - (I \cup J)) \notin \mathcal{K}$. Now, (t) leads to $(i,j|N - (I \cup J)) \notin \mathcal{K}$ what implies the existence of a set $M_{i,j} \in \mathcal{K}^*$ containing $i \cup j$ and contained in $I \cup J$. Finally observe that the family $\{I, J\} \cup \{M_{i,j};\ i \in I - J,\ j \in J - I\} \subset \mathcal{K}^*$ realizes a double union of $I \cup J$ and thus $I \cup J \in \mathcal{K}^*$. We obtained $\mathcal{K}^* \in \mathbf{V}(N)$.

Let us suppose $(i,k|K) \notin \mathcal{K},\ (j,k|K) \notin \mathcal{K}$, where $\mathcal{K} \in \mathbf{A}(N)$ and $\mathcal{K}^* \in \mathbf{V}(N)$. By Theorem 1 there exist $L, M \in \mathcal{K}^*$ such that $i \cup k \subset L,\ j \cup k \subset M$ and $(L \cup M) \cap K = \emptyset$. Since L, M have nonempty intersection $L \cup M \in \mathcal{K}^*$. Knowing that $(i,j|K) \,ø\, (L \cup M)$ Theorem 1 applied again yields $(i,j|K) \notin \mathcal{K}$. □

We can conclude that there is a bijection between the classes of all I–relations and C–families. This fact was proved directly in Matúš (1988) by the Galois correspondence methods.

9 REMARKS

Definition 9. *A subset $\mathcal{L}$ of $\mathcal{Q}(N)$ is called p–representable (dually p–representable) if there exists ξ such that*

$$(I,J) \in \mathcal{L} \iff \xi_I \perp \xi_J \,|\, \xi_\emptyset \qquad \left(\xi_I \perp \xi_J \,|\, \xi_{N-(I\cup J)}\right).$$

Any p–representable $\mathcal{L} \subset \mathcal{Q}(N)$ is an I–relation on N as we see from (CI 0) and (CI 1). On the contrary, any I–relation $\mathcal{L}$ can be written in the form

$$\mathcal{L} = \left\{(I,J) \in \mathcal{Q}(N);\ (I,J|\emptyset) \in \mathrm{gl}(\mathrm{con}(\mathcal{L}))\right\}.$$

Since $\mathrm{gl}(\mathrm{con}(\mathcal{L}))$ is a global DCI–relation, it is p–representable due to Theorem 5 and so is $\mathcal{L}$ too. In Matúš (1988) or Matúš (1990 b) one can find a direct construction of a p–representation of any I–relation by a system of two–valued random variables.

The case of dual p–representations is dealt with similarly. If $\mathcal{L}$ is dually p–representable then (CI 0) and (CI 1) cause it is an I–relation. If $\mathcal{L} \in \mathbf{I}(N)$ then (cf. Proposition 3)

$$\mathcal{L} = \left\{(I,J) \in \mathcal{Q}(N);\ (I,J|N-(I\cup J)) \in \mathrm{gl}(\mathrm{con}(\mathcal{L}))^\rceil\right\}$$

and since $\mathrm{gl}(\mathrm{con}(\mathcal{L}))^\rceil$ is a global ACI–relation it is p–representable and thus $\mathcal{L}$ is dually p–representable.

We proceed by a discussion of the subclass of I–relations determined by

(G) $$(I,J) \in \mathcal{L},\ (I,K) \in \mathcal{L} \implies (I,J\cup K) \in \mathcal{L}.$$

Every I–relation of this sort is uniquely given by the simple graph $G = (N,H)$, where ij is an edge of H if $(i,j) \in \mathcal{L}$, and vice versa.

The embedding con maps these graphical I–relations onto the subclass of $\mathbf{D}(N)$ specified by (b) and by

(g) $$(i,j|L) \in \mathcal{K},\ (i,k|L) \in \mathcal{K} \implies (i,j|k\cup L) \in \mathcal{K}$$

(or equivalently by (d), (b) and (g), since (d) (g) $\Rightarrow$ (d1)). Notice that $(i,j|K) \in \mathrm{con}(\mathcal{L})$ just if there is no path between vertices i and j passing through K.

Analogically, the embedding $\mathcal{L} \to (\mathrm{con}(\mathcal{L}))^\rceil$ maps the graphical I–relations onto the subclass of $\mathbf{A}(N)$ specified by (t) and by the dual property of (g)

(m) $$(i,j|k\cup L) \in \mathcal{K},\ (i,k|j\cup L) \in \mathcal{K} \implies (i,j|\,L) \in \mathcal{K}$$

(or equivalently by (a), (t) and (m)). Notice that $(i,j|K) \in (\mathrm{con}(\mathcal{L}))^\rceil$ just if K is a cutset of i and j, so that we arrived at the conditional independence structures of Markov networks (cf. Pearl (1986)). It is easy to check that the C–family $(\mathrm{con}\,(\mathcal{L}))^{\rceil *}$ consists of the connected sets of the graph G.

Acknowledgement

I record my gratitude to Mr. M. Studený for many challenging discussions during the last year and for a constructive criticism of this work.

REFERENCES

Birkhoff G. (1967) : Lattice Theory. Providence R. I.

Dawid A. P. (1979) : Conditional independence in statistical theory (with discussion). Journal of the Royal Stat. Soc. B 41, 1–31.

Geiger D., Paz A. and Pearl J. (1989) : Axioms and algorithms for inferences involving probabilistic independence. UCLA, Cognitive System Lab., Technical Report R–119.

Geiger D., Paz A. and Pearl J. (1989) : Logical and algorithmic properties of independence and their application to Bayesian networks. UCLA, Cognitive System Lab., Technical Report R–123.

Hausdorff F. (1914) : Grundzüge der Mengenlehre. Lepzig.

Malvestuto F. M. (1990) : A unique formal system for binary decompositions of database relations, probability distributions and graphs. Information Sciences (to appear).

Martin N. F. G. and England J. W. (1981) : Mathematical Theory of Entropy. Addison–Wesley Publ. Comp., London.

Matúš F. (1988) : Nezávislosť a Radonove projekcie na kompaktných grupách (Independence and Radon projections on compact groups). Ph. D. Thesis, Prague.

Matúš F. (1990 a) : Abstract functional dependence structures. (submitted to) Theor. Comp. Science.

Matúš F. (1990 b) : Independence structures modelling stochastic and algebraic independences and connectedness. (submitted to) Informaton and Computation.

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

Neveu J. (1964) : Bases Mathématiques du Calcul des Probabilités. Mason et C^{ie}, Paris.

Paredaens J. (1980) : The interaction of integrity constraints in an information system. Journal of Comp. and System Sciences 20, 310–329.

Pearl J. (1986) : Markov and Bayes networks: a comparision of two graphical representations of probabilistic knowledge. UCLA, Cognitive Syst. Lab., Technical report R–46.

Studený M. (1989) : Multiinformation and the problem of characterization of conditional independence relations. Problems of Control and Inf. Th. 18, 3–16.

Studený M. (1990) : Conditional independence relations have no finite complete characterization. (this proceedings)

Welsh D. J. A. (1976) : Matroid Theory. Academic Press, London.

White N. ed. (1986) : Theory of Matroids. Cambridge University Press, Cambridge.

Institute of Information Theory and Automation
Czechoslovak Academy of Sciences
Pod vodárenskou věží 4
182 08 Prague 8
Czechoslovakia