CONDITIONAL INDEPENDENCE RELATIONS HAVE NO FINITE COMPLETE CHARACTERIZATION

Milan Studený

Prague

Key words: Conditional–independence relation, Axiomatization of conditional independence, Syntactic description of conditional independence, Embedded multivalued dependency

ABSTRACT

The hypothesis of existence of a finite characterization of conditional–independence relations (CIRs) is refused. This result is shown to be equivalent with the non–existence of a simple deductive system describing relationships among CI–statements (it is certain type of syntactic description). However, under the assumption that CIRs are grasped the existence of a countable characterization of CIRs is shown. Finally, the problem of characterization of CIRs is shown to be diverse from an analogical problem of axiomatization EMVDs arising in the theory of relational databases.

INTRODUCTION

Let $[\xi_i]_{i \in N}$ be a random vector ($2 \leq \operatorname{card} N < \infty$) and let us suppose for simplicity that its components are finite–valued random variables. Then we can define a ternary disjoint relation I on $\exp N$ (disjoint means that its domain is the set of triplets of pairwise disjoint subsets of N):

$I(A;\, B|C)$ holds iff $[\xi_i]_{i \in A}$ is conditionally independent of $[\xi_i]_{i \in B}$ given $[\xi_i]_{i \in C}$.

We shall call this relation the *conditional–independence relation* (CIR) corresponding to $[\xi_i]_{i \in N}$ as it describes all conditional–independence relationships among its subvectors.

Our question is whether it is possible to characterize CIRs as ternary disjoint relations satisfying a set of properties of the following type:

$$\square \qquad [I(A_1, B_1|C_1) \ \&\ I(A_2, B_2|C_2) \ \&\ \cdots \ \&\ I(A_r, B_r|C_r)] \ \longrightarrow I(A_{r+1}, B_{r+1}|C_{r+1})$$

This task was called *problem of characterization of* CIR*s* in [16]. Many authors dealing with related problems ([4], [6]) speak about *axiomatization of conditional independence* (CI). Indeed, finding a finite set of such properties characterizing CIRs it is possible to describe relationships among CI-statements syntactically in simple way (i. e. by means of a deductive system – this notion paraphrases the notion of formal axiomatic theory from mathematical logic). We shall discuss it in §5.

Nevertheless, in this article we prove that CIRs cannot be characterized by a finite set of properties of the type $\square$. More exactly, we shall find for every $4 \leq n = \text{card } N$ some property of CIRs (of the type $\Box$) which cannot reveal in lower dimensions. Then we shall use this result to show that no deductive system whose formulas correspond to individual CI–statements (we speak about *simple* deductive system for describing CI) can completely comprehend relationships among CI–statements.

The situation has analogy in the theory of relational databases where *embedded multivalued dependencies* (EMVDs) were shown have no complete axiomatization [15]. One can formulate a hypothesis that CIRs and complexes of EMVDs can be characterized by the same set of properties of the type $\Box$. We give two examples to refuse it. The former $\Box$–property holds for CIRs but fails in case of EMVD, the latter one holds in case of EMVD but does not hold for CIRs.

The first section deals with sources and history of the problem and gives the corresponding references. The second section gives basic definitions, the third one contains some preparatory results. The main result is proved in the fourth section. It is supplied by a construction of a countable characterization of CIRs by $\Box$–properties (however applicable only in case that all CIRs are grasped). The fifth section is devoted to syntactical description of CI. We discuss how a formal axiomatization for CI can look and show that no simple complete deductive system for CI exists. The sixth section analyses the analogy with the theory of relational databases and gives the promised examples. In the last section (concluding remarks) we summarize the article and propose a plan of further investigation.

1 HISTORY OF THE PROBLEM

The conditional independence (CI) is one of the basic concepts of probability theory. Its importance in modern statistics was accentuated by Dawid [2] twelve years ago, where some formal properties of CI was noticed. Since that time many articles have been concerned with those properties (for example [10], [14]).

Our interest in this problem is motivated by its expected profit in the theory of probabilistic expert systems. The notion of CI can be interpreted as certain (nonnumerical) relationship among symptoms (which are described by random variables) and thus it promises the possibility to determine the proper structure of the expert system directly by asking experts (see [16], [13]).

The importance of CI for probabilistic expert systems was explicitly discerned and highlighted especially by Pearl who in [12] formulated a concrete conjecture for characterization of CIRs corresponding to random vectors whose distributions are strictly positive measures. This conjecture was refused in [16] by finding a new independent property of CIRs. Note that another task motivated by the same work was solved in [5].

Nevertheless, some positive results were achieved in this respect, namely certain subclasses of CI–statements were characterized. In [4] and independently in [7] a complete characterization for the class of "marginal" CI–statements (i. e. statements $I(A;\ B|C)$ with fixed C) was found. In [3] and also in [6] (using different formal description) the class of "fixed–context" CI–statements (i. e. $I(A;B|C)$ where $A \cup B \cup C$ is fixed) was characterized. In connection with these result we would like to bring readers attention to the article [8] where special classes of CIRs ("monotonic" in condition) are characterized.

The situation is similar to the situation in the theory of relational databases where attempts to axiomatize miscellaneous types of dependencies were made. Relationships among multivalued dependencies (MVDs) were completely characterized in [1]. We can consider certain analogy, where the axiomatization of MVDs corresponds with the characterization for the class of "fixed–context" CI–statement, and axiomatization of embedded multivalued dependencies (EMVD) corresponds with the problem of characterization of CIRs. Thus, this article gives a result quite analogical to the result from [15] saying that EMVDs have no complete axiomatization. Indeed, we found inspiration in [15].

On the other hand the analogy is not absolute as the reader can see from the examples in the sixth section.

2 BASIC DEFINITIONS

In all this article *index set* will be any finite set having at least two elements. Given an index set N let us denote by $T(N)$ the set of all ordered triplets of pairwise disjoint subsets of N. Every subset of $T(N)$ will be called *dependency model on* N (terminology borrowed from [3]). Given $u = \langle A, B, C\rangle \in T(N)$ its *context* is the set $A \cup B \cup C$. It will be denoted by $[u]$. We shall say that $u = \langle A, B, C\rangle \in T(N)$ is *trivial* iff A or B is empty. The set of non–trivial triplets will be denoted by $T_*(N)$.

Let K and L be two index sets and $v : K \to L$ an injective mapping (necessarily card $K \leq$ card L). Then it can be considered as an injective mapping $\boldsymbol{v} : T(K) \to T(L)$:

$$\langle A, B, C\rangle \mapsto^{\boldsymbol{v}} \langle v(A), v(B), v(C)\rangle \qquad \text{whenever} \qquad \langle A, B, C\rangle \in T(K).$$

Moreover, for every positive integer $r \geq 1$ it can be considered as an injection $\overline{\boldsymbol{v}} : T(K)^{r+1} \to T(L)^{r+1}$:

$$(u_1, \ldots, u_{r+1}) \mapsto^{\overline{\boldsymbol{v}}} (\boldsymbol{v}(u_1), \ldots, \boldsymbol{v}(u_{r+1})) \qquad \text{whenever} \qquad u_1, \ldots, u_{r+1} \in T(K).$$

For every index set N we introduce the class of N–dimensional measures $\mathcal{P}(N)$. Every element of $\mathcal{P}(N)$ is specified by a collection of finite nonempty sets $\{X_i;\ i \in N\}$ and by a probability measure on $\prod_{i\in N} X_i$ (endowed with the σ–algebra of all subsets). Another possible view: elements of $\mathcal{P}(N)$ are distributions of N–dimensional random vectors $[\xi_i]_{i\in N}$ where ξ_i takes values in X_i.

Having an index set N and $P \in \mathcal{P}(N)$ we define a dependency model I on N as follows:

- if $\langle A, B, C\rangle \in T_*(N)$ and $C \neq \emptyset$ then $\quad \langle A, B, C\rangle \in I \quad$ iff $\quad P^{A\cup B\cup C} \cdot P^C = P^{A\cup C} \cdot P^{B\cup C}$
  (P^A denotes the marginal measure of P on $\prod_{i\in A} X_i$)
- if $\langle A, B, C\rangle \in T_*(N)$ and $C = \emptyset$ then $\quad \langle A, B, \emptyset\rangle \in I \quad$ iff $\quad P^{A\cup B} = P^A \cdot P^B$
- if $\langle A, B, C\rangle \in T(N) \setminus T_*(N)$ then we postulate $\langle A, B, C\rangle \in I$

We shall call this dependency model the *conditional–independence relation* (abbreviation CIR) corresponding to P. The symbol $CIR(N)$ will be used to denote the class of CIRs corresponding to all measures from $\mathcal{P}(N)$. We shall sometimes use the symbol $I(A;\ B|C)$ instead of $\langle A, B, C\rangle \in I$.

Informally, by $\Box$–rule we shall understand the schema:

$\Box$
$$[I(A_1, B_1|C_1)\ \&\ \cdots\ \&\ I(A_r, B_r|C_r)] \longrightarrow I(A_{r+1}, B_{r+1}|C_{r+1})$$

The triplets $I(A_1, B_1|C_1), \ldots, I(A_r, B_r|C_r)$ are called *antecedents*, the triplet $I(A_{r+1}, B_{r+1}|C_{r+1})$ the *consequent*. Of course, we are supposed to be able to substitute concrete elements of $T(N)$ for $I(A_i;\ B_i|C_i)$ (for every index set N). Certainly, this notion requires a precise definition. Thus, a $\Box$–rule is specified by a nonnegative integer r and by a collection of $(r+1)$–nary relations on $T(N)$: $\{\boldsymbol{A}(N) \subset T(N)^{r+1};\ N \text{ is an index set}\}$
For example, take $r = 1$ and put (for every N):

$$\boldsymbol{A}_{\bullet}(N) = \{[\langle A, B, C\rangle, \langle B, A, C\rangle];\ A, B, C \subset N \text{ are pairwise disjoint}\}.$$

This $\Box$–rule we easily express by the schema $[I(A; B|C)] \longrightarrow I(B; A|C)$.

Another examples of $\Box$–rules are described here:

$$[I(A; B|C \cup D)\ \&\ I(A; C|D)] \longrightarrow I(A; B \cup C|D) \qquad [\,] \longrightarrow I(\emptyset; B|C).$$

All these □–rules meet the following special property. A □–rule (determined by $\boldsymbol{A}(N) \subset T(N)^{r+1}$ for any N) is *regular* iff

$$\forall\ K, L \text{ index set} \quad v: K \to L \text{ injective} \quad \boldsymbol{A}(K) = (\overline{\boldsymbol{v}}_{-1})\, \boldsymbol{A}(L). \tag{2.1}$$

Given an index set N and a □–rule we shall say that a dependency model $I \subset T(N)$ is *closed* under this □–rule iff it holds: whenever $u_1, \ldots, u_{r+1}$ can be substituted for antecedents and consequent of this □–rule, then $\{u_1, \ldots, u_r\} \subset I$ implies $u_{r+1} \in I$ (exactly $u_1, \ldots, u_r \in I$ & $(u_1, \ldots, u_{r+1}) \in \boldsymbol{A}(N)$ implies $u_{r+1} \in I$). It seems to us that rule is more adequate term that axiom used in [4] and [6] (see § 5 Remark 2).

In the sequel we shall often meet the index set $N = \{0, 1, \ldots, n\}$ where $n \geq 3$. In that situation we shall consider the following operation of *successor* $\mathrm{suc} : \{1, \ldots, n\} \to \{1, \ldots, n\}$:
$\mathrm{suc}\,(i) = i + 1$ whenever $i = 1, \ldots, n-1$ $\qquad \mathrm{suc}\,(n) = 1$.

3 AUXILIARY RESULTS CONCERNING CIRs

Firstly, we mention a trivial property of CIRs namely *symmetry.* Given an index set N and $I \in CIR(N)$ it holds:

$$\langle A, B, C\rangle \in I \quad \text{iff} \quad \langle B, A, C\rangle \in I \quad \text{whenever} \quad \langle A, B, C\rangle \in T(N) \tag{3.1}$$

In the sequel we construct some concrete examples of CIRs. But constructions of probability measures are given only. The verification that the corresponding CIRs meet our requirements is left to the reader. Note that it can be easily done using techniques from [18].

The first lemma contains a construction borrowed from [3] which simplifies our tasks.

Lemma 1. *Let N be an index set and $I, J \in CIR(N)$. Then $I \cap J \in CIR(N)$ (as $I, J \subset T(N)$ also $I \cap J \subset T(N)$).*

Proof. Suppose that $P \in \mathcal{P}(N)$ (on $\prod_{i \in N} X_i$) gives I as its corresponding CIR and $Q \in \mathcal{P}(N)$ (on $\prod_{i \in N} Y_i$) gives J. Define $R \in \mathcal{P}(N)$ on $\prod_{i \in N} X_i \times Y_i$ as follows:

$$R\left([x_i,\, y_i]_{i \in N}\right) = P\left([x_i]_{i \in N}\right) \cdot Q\left([y_i]_{i \in N}\right) \qquad \text{where} \qquad x_i \in X_i,\ y_i \in Y_i$$

It can be verified that $I \cap J$ is the CIR corresponding to R. △

Lemma 2. *Supposing $N = \{0, 1, \ldots, n\}$ $(n \geq 3)$ and $s \in \{1, \ldots, n\}$ we put:*

$$I = \bigcup_{j \in \{1, \ldots, n\} \setminus \{s\}} \{\langle 0, j, suc(j)\rangle, \langle j, 0, suc(j)\rangle\} \cup [T(N) \setminus T_*(N)]\,.$$

Then $I \in CIR(N)$.

Proof. Firstly we give three auxiliary constructions.

I Having $D \subset N$, $\mathrm{card}\, D \geq 2$ it holds
$I_D = \{\langle A, B, C\rangle \in T(N);\ A \cap D = \emptyset \text{ or } B \cap D = \emptyset \text{ or } D \not\subset A \cup B \cup C\} \in CIR(N)$.

Indeed: Put $X_i = \{0, 1\}$ for $i \in D$ and $X_i = \{0\}$ for $i \in N \setminus D$ and define a distribution P on $\prod_{i \in N} X_i$ as follows:

$$P\left([x_i]_{i \in N}\right) = \begin{cases} 2^{1 - \mathrm{card}\, D} & \text{if } \sum_{i \in D} x_i \text{ is even} \\ 0 & \text{if } \sum_{i \in D} x_i \text{ is odd.} \end{cases} \qquad \otimes$$

II There exists $K \in CIR(N)$ satisfying

- $\langle 0, i, j\rangle \in K$ whenever $i, j \in \{1, \ldots, n\}$ $i \neq j$
- $\langle i, j, 0\rangle \notin K$ whenever $i,$ $j \in \{1, \ldots, n\}$ $i \neq j$
- $\langle A, B, \emptyset\rangle \notin K$ whenever $A,$ $B \neq \emptyset$.

Indeed: Define $P \in \mathcal{P}(N)$ on $\prod_{i \in N} X_i$ where $X_i = \{0, 1\}$ (for each $i \in N$) as follows:

$$\begin{array}{ll} (0,\ 0,\ 0,\ \ldots, 0) & \longrightarrow a_1 \\ (1,\ 0,\ 0,\ \ldots, 0) & \longrightarrow a_2 \\ (0,\ 1,\ 1,\ \ldots, 1) & \longrightarrow a_3 \\ (1,\ 1,\ 1,\ \ldots, 1) & \longrightarrow a_4 \end{array}$$

where $a_1, \ldots, a_4 > 0,\ a_1 + \cdots + a_4 = 1,\ a_1 \cdot a_4 \neq a_2 \cdot a_3$. $\otimes$

III There exists $J \in CIR(N)$ satisfying

- $\langle 0, j, \mathrm{suc}(j)\rangle \in J$ whenever $j = 1, \ldots, n-1$
- $\langle 0, \mathrm{suc}\,(j), j\rangle \notin J$ whenever $j = 1, \ldots, n-1$.

Indeed: Put $X_i = \{1, \ldots, i\}$ for $i \in \{1, \ldots, n\}$ and $X_0 = X_n$. Define $a^1, \ldots, a^n \in \prod_{i \in N} X_i$ as follows: $a^k = \left(a_0^k, a_1^k, \ldots, a_n^k\right)$ where $a_0^k = k$ and $a_j^k = \min\{k, j\}$ whenever $j = 1, \ldots, n$. Consider the uniform distribution on $\left\{a^1, \ldots, a^n\right\}$ i.e. $a^k \to n^{-1}$ for every $k \in \{1, \ldots, n\}$. $\otimes$

IV Clearly, without loss of generality we can suppose $s = n$. Thus, we denote

$$I = \bigcup_{j=1}^{n-1} \{\langle 0, j, \mathrm{suc}\,(j)\rangle,\ \langle j, 0, \mathrm{suc}\,(j)\rangle\} \ \cup\ [T(N) \setminus T_*(N)]\,.$$

Our task is to show $I \in CIR(N)$. For this purpose we put $\mathcal{D} = \mathcal{D}_1 \cup \mathcal{D}_2$ where

$$\begin{array}{rcl} \mathcal{D}_1 & = & \{D \subset N;\ \mathrm{card}\, D = 4\} \\ \mathcal{D}_2 & = & \{D \subset N;\ \mathrm{card}\, D = 3\ \&\ D \neq \{0, j, \mathrm{suc}\,(j)\}\ \text{for every}\ j = 1, \ldots, n-1\}\,. \end{array}$$

Consider dependency models I_D for $D \in \mathcal{D}$ (see I), K from II and J from III. By Lemma 1 $K \cap J \cap \bigcap_{D \in \mathcal{D}} I_D \in CIR(N)$. Evidently $I \subset K \cap J \cap \bigcap_{D \in \mathcal{D}} I_D$. To verify the reversed inclusion consider $u = \langle A, B, C\rangle \in T(N) \setminus I$ and show $u \notin K \cap J \cap \bigcap_{D \in \mathcal{D}} I_D$:

- if $\mathrm{card}\,[u] \geq 4$ then $u \notin \bigcap_{D \in \mathcal{D}_1} I_D$ (see I)
- if $C = \emptyset$ (special case $\mathrm{card}\,[u] = 2$) then $u \notin K$ (see II)
- $\mathrm{card}\,[u] = 3$ & $[u] \in \mathcal{D}_2$ then $u \notin \bigcap_{D \in \mathcal{D}_2} I_D$ (see I)
  Thus, consider $\mathrm{card}\,[u] = 3$ & $[u] \notin \mathcal{D}_2$ & $C \neq \emptyset$. Then $[u] = \{0, j, \mathrm{suc}\,(j)\}$ for some $j \in \{1, \ldots, n-1\}$ and only two cases can occur:
- $u \in \{\langle j, \mathrm{suc}(j), 0\rangle,\ \langle \mathrm{suc}(j), j, 0\rangle\} \ldots$ then $u \notin K$ (see II)
- $u \in \{\langle 0, \mathrm{suc}(j), j\rangle,\ \langle \mathrm{suc}(j), 0, j\rangle\} \ldots$ then $u \notin J$ (see III) $\triangle$

Lemma 3. *Let K, L be index sets, $v : K \to L$ an injective mapping.*

a) *Given $I \in CIR(K)$ there exists $J \in CIR(L)$ such that*

(3.2) $$u \in I \Longleftrightarrow \boldsymbol{v}(u) \in J \qquad \text{whenever} \qquad u \in T(K)$$

b) *Given $J \in CIR(L)$ there exists $I \in CIR(K)$ such that* (3.2) *holds.*

Proof.

a) Let $P \in \mathcal{P}(K)$ (on $\prod_{i \in K} X_i$) gives I as its corresponding CIR. Put $Y_j = X_{v_{-1}(j)}$ for $j \in v(K)$ and $Y_j = \{0\}$ for $j \in L \setminus v(K)$. The natural injection of $\prod_{i \in K} X_i$ into $\prod_{j \in L} Y_j$ is a one–to–one mapping and transfers P to a measure Q on $\prod_{j \in L} Y_j$. Take J as the CIR corresponding to $Q \in \mathcal{P}(L)$.

b) Let $Q \in \mathcal{P}(L)$ (on $\prod_{j \in L} Y_j$) gives $J \in CIR(L)$. Put $X_i = Y_{v(i)}$ for $i \in K$. The natural bijection between $\prod_{i \in K} X_i$ and $\prod_{j \in v(K)} Y_j$ transfers the marginal of Q on $\prod_{j \in v(K)} Y_j$ to some prob. measure P on $\prod_{i \in K} X_i$. Define I as its corresponding CIR. $\triangle$

Lemma 4. *Let N be an index set, $V \subset N$ and $u \in T_*(N)$ with $[u] \setminus V \neq \emptyset$. Then there exists $I \in CIR(N)$ such that $u \notin I$ and*

$$w \in I \qquad \textit{whenever} \qquad w \in T(N), [w] \subset V.$$

Proof. Put $D = [u]$ and use part I of the proof of Lemma 2. $\triangle$

4 MAIN RESULTS

A new property of CIRs is derived in the following proposition.

Proposition 1. *Let $K = \{0, \dots, k\}$ where $k \geq 3$. Consider the operation of successor on $\{1, \dots, k\}$. Let $I \in CIR(N)$ where $K \subset N$. Then the following conditions are equivalent:*

(a) $\forall\, j = 1, \dots, k \qquad I\,(0;\, j|\mathrm{suc}\,(j))$

(b) $\forall\, j = 1, \dots, k \qquad I\,(0;\, \mathrm{suc}\,(j)|j)$

Proof. We shall use the following result from [16]: We can assign the multiinformation function $I_m :\ \exp N \rightarrow \langle 0, \infty)$ satisfying the following conditions (4.1), (4.2) to every $P \in \mathcal{P}(N)$:

(4.1) $$I_m[A \cup B \cup C] + I_m[C] \geq I_m[A \cup C] + I_m[B \cup C] \text{ for every } \langle A, B, C\rangle\ \in T(N)$$

(4.2) $I(A; B|C)$ holds iff the equality in (4.1) holds (I denotes the CIR corresponding to P).

The condition (a) can be written by (4.2) as follows:

$\forall\, j = 1, \dots, k \quad I_m[0, j,\, \mathrm{suc}\,(j)] + I_m[\mathrm{suc}\,(j)] - I_m[0,\, \mathrm{suc}\,(j)] - I_m[j,\, \mathrm{suc}\,(j)] = 0$

Hence, we get:

$$\begin{aligned}
0 &= \sum_{j=1}^{k} \{\ I_m[0, j,\, \mathrm{suc}\,(j)] + I_m[\mathrm{suc}\,(j)] - I_m[0, \mathrm{suc}\,(j)] - I_m[j, \mathrm{suc}\,(j)]\ \} = \\
&= \sum_{j=1}^{k} I_m[0, j,\, \mathrm{suc}\,(j)] + \sum_{j=1}^{k} I_m[\mathrm{suc}\,(j)] - \sum_{j=1}^{k} I_m[0,\, \mathrm{suc}\,(j)] - \sum_{j=1}^{k} I_m[j,\, \mathrm{suc}\,(j)] = \\
&= \sum_{j=1}^{k} I_m[0, j,\, \mathrm{suc}\,(j)] + \sum_{j=1}^{k} I_m[j] - \sum_{j=1}^{k} I_m[0, j] - \sum_{j=1}^{k} I_m[j,\, \mathrm{suc}\,(j)] = \\
&= \sum_{j=1}^{k} \{\ I_m[0, j,\, \mathrm{suc}\,(j)] + I_m[j] - I_m[0, j] - I_m[j,\, \mathrm{suc}\,(j)]\ \}
\end{aligned}$$

According to (4.1) every expression in braces in the last sum is nonnegative. Thus, it vanishes and using (4.2) we easily derive (b).
The implication (b) $\Rightarrow$ (a) can be verified analogously. $\triangle$

Hence, we derive our main result.

Consequence 1. *No finite system $(\mathcal{A}.1) - (\mathcal{A}.p)$ of $\square$–rules can characterize CIRs in sense that for every dependency model I on any index set N it holds*
$[I \in \ CIR(N) \quad$ *iff* $\quad I$ *is closed under* $(\mathcal{A}.1) - (\mathcal{A}.p)]$.

Proof: Let $(\mathcal{A}.1), \ldots, (\mathcal{A}.p)$ be a finite set of $\square$–rules. Suppose that every CIR is closed under $(\mathcal{A}.1) - (\mathcal{A}.p)$. It suffices to find a dependency model closed under $(\mathcal{A}.1) - (\mathcal{A}.p)$ which is not CIR for any probability measure.
Let $n \geq 3$ exceeds the maximal number of antecedents of $(\mathcal{A}.i)$ $(i = 1, \ldots, p)$. Put $N = \{0, \ldots, n\}$, consider the operation of successor on N and define $I \subset T(N)$ as follows:

$$I = \bigcup_{j=1}^{n} \{\langle 0,\, j,\, \mathrm{suc}\,(j)\rangle,\ \langle j,\, 0,\, \mathrm{suc}\,(j)\rangle\} \cup [T(N) \setminus T_*(N)]\,.$$

To show that I is closed under $(\mathcal{A}.i)$ $(i = 1, \ldots, p)$ consider a set $K \subset I$ that can be substituted for antecedents of $(\mathcal{A}.i)$. Let $u \in T(N)$ can be substituted for the corresponding consequent. As $K \subset I$ and $\mathrm{card}\, K < n$, according to Lemma 2 we can find $J \in CIR(N)$ such that $K \subset J \subset I$. Certainly, J is closed under $(\mathcal{A}.i)$. Hence $K \subset J$ gives $u \in J \subset I$. Thus, I is closed under $(\mathcal{A}.i)$. Nevertheless, according to Proposition 1 $\quad I \notin CIR(N)$. $\triangle$

Although CIRs cannot be characterized by finite number of $\square$–rules they can be characterized by countably many regular $\square$–rules provided we can grasp CIRs. The construction is contained in the following proposition.

Proposition 2. *Let us suppose that we have a list of elements $CIR(N)$ for any index set N at our disposal. Then we can construct a countable set of regular $\square$–rules $(\mathcal{A}.r)$ $r = 0, 1, \ldots$ satisfying the conditions a) – d):*

a) *every $(\mathcal{A}.r)$ has exactly r antecedents i. e. it is given by $\boldsymbol{A}_r(N) \subset T(N)^{r+1}$*

b) $\boldsymbol{A}_0(N) = T(N) \setminus T_*(N)$ $\quad$ (*$(\mathcal{A}.0)$ includes exactly trivial triplets)*

c) $\boldsymbol{A}_r(N) \subset T_*(N)^{r+1}$ $\qquad$ *whenever $r \geq 1$*

d) *every $(\mathcal{A}.r)$ for $r \geq 1$ respects the context i. e.*
$\quad (u_1, \ldots, u_{r+1}) \in \boldsymbol{A}_r(N) \quad$ *implies* $\quad [u_{r+1}] \subset \bigcup_{i=1}^{r} [u_i]$

such that for every dependency model I on any index set N it holds:
$[I \in \ CIR(N) \quad$ *iff* $\quad I$ *is closed under all* $(\mathcal{A}.r)$ $r = 0, 1, \ldots]$.

Proof. Given $r = 0, 1, \ldots$, an index set N, $u_1, \ldots, u_{r+1} \in T(N)$ we define:

$$(u_1, \ldots, u_{r+1}) \in \boldsymbol{A}_r(N) \iff \begin{array}{ll} (1) & \forall\, I \in CIR(N)\ u_1, \ldots, u_r \in I \quad \text{implies} \quad u_{r+1} \in I \\ (2) & \forall\, S \subsetneq \{u_1, \ldots, u_r\}\ \exists\, I \in CIR(N) \quad S \subset I \ \&\ u_{r+1} \notin I \\ (3) & u_1, \ldots, u_r \text{ are distinct} \end{array}$$

Thus, the set of $\square$–rules $(\mathcal{A}.r)$ $r = 0, 1, \ldots$ is introduced.

I Given $r = 0, 1, \ldots$, $v : K \to L$ injective it holds $\boldsymbol{A}_r(K) = (\overline{\boldsymbol{v}}_{-1})\, \boldsymbol{A}_r(L)$ (i.e. $(\mathcal{A}.r)$ is regular)
Indeed: Equivalently, given $u_1, \ldots, u_{r+1} \in T(K)$ we need to show

$$(u_1, \ldots, u_{r+1}) \in \boldsymbol{A}_r(K) \qquad \text{iff} \qquad (\boldsymbol{v}(u_1), \ldots, \boldsymbol{v}(u_{r+1})) \in \boldsymbol{A}_r(L)$$

It makes no problem verify it using Lemma 3. $\otimes$

II $I \in CIR(N) \Rightarrow I$ is closed under $(\mathcal{A}.r)$ $r = 0, 1, \ldots$
Indeed: having $(u_1, \ldots, u_{r+1}) \in \boldsymbol{A}_r(N)$ $u_1, \ldots, u_r \in I$, by the condition (1) we derive $u_{r+1} \in I$. $\otimes$

III Let $I \subset T(N)$ be closed under $(\mathcal{A}.r)$ $r = 0, 1, \ldots$. Then $I \in CIR(N)$.
Indeed: Owing to Lemma 1 it suffices to show: for every $u \in T(N) \setminus I$ there exists $J \in CIR(N)$ with $I \subset J$ and $u \notin J$ (note that $T(N)$ is finite). Equivalently, for every $u \in T(N) \setminus I$ the system $\mathcal{V}_u = \{K \subset I;\ \forall\, J \in CIR(N)\ K \subset J \text{ implies } u \in J\}$ is empty. In case $\mathcal{V}_u \neq \emptyset$ we take some its minimal element $K_0 = \{v_1, \ldots, v_k\}$. It is easy to verify $(v_1, \ldots, v_k, u) \in \boldsymbol{A}_k(N)$. Hence, $u \in I$ (as I is closed under $(\mathcal{A}.k)$) and it contradicts the original selection of u. $\otimes$

IV $\boldsymbol{A}_0(N) = \bigcap_{I \in CIR(N)} I = T(N) \setminus T_*(N)$
Indeed: The first equality is evident. Having $u \in T(N) \setminus T_*(N)$ the definition of CIRs gives $u \in \bigcap_{I \in CIR(N)} I$. On the other hand $u \in T_*(N)$ implies $u \notin \bigcap_{I \in CIR(N)} I$ by Lemma 4 (with $V = \emptyset$). $\otimes$

V $\boldsymbol{A}_r(N) \subset T_*(N)^{r+1}$ whenever $r \geq 1$
Indeed: Let $(u_1, \ldots, u_{r+1}) \in \boldsymbol{A}_r(N)$. By (2) for $S = \emptyset$ we get $u_{r+1} \notin \bigcap_{I \in CIR(N)} I \equiv T(N) \setminus T_*(N)$ (see IV). Suppose that $u_j \notin T_*(N)$ for some $j \leq r$ and put $S = \{u_1, \ldots, u_r\} \setminus \{u_j\}$. By (3), (2) we find $I \in CIR(N)$ such that $S \subset I$ and $u_{r+1} \notin I$. By IV $u_j \in I$ i.e. $\{u_1, \ldots, u_r\} \subset I$. It contradicts (1). $\otimes$

VI $u_1, \ldots, u_{r+1} \in T(N)$ $r \geq 1$ $(u_1, \ldots, u_{r+1}) \in \boldsymbol{A}_r(N)$ implies $[u_{r+1}] \subset \bigcup_{i=1}^{r} [u_i]$
Indeed: Supposing $[u_{r+1}] \setminus \bigcup_{i=1}^{r} [u_i] \neq \emptyset$ we put $V = \bigcup_{i=1}^{r} [u_i]$ and $u = u_{r+1}$ and apply Lemma 4 to find $I \in CIR(N)$ with $u_1, \ldots, u_r \in I$ and $u_{r+1} \notin I$. It contradicts the condition (1). $\triangle$

5 SYNTACTIC DESCRIPTION OF CI

Firstly, we are going to introduce the notion of *deductive system* to be a tool for syntactic description instead of the notion of formal axiomatic theory (see [9], chap. 1, sec. 4). Actually, these notions are almost coincident but the notion of deductive system is more general. We omit the principal requirement claimed on a formal axiomatic theory that is recursivity (i.e. effective determination of formulas, axioms and inference rules).

Definition 1 (deductive system).
Deductive system D is defined when the following conditions are satisfied:

(1) *A countable set of symbols S is given. Finite sequences of symbols are called expressions of D.*

(2) *There is a subset F of the set of expressions called the set of formulas of D.*

(3) *A set of formulas A is set aside and called the set of axioms of D.*

(4) *There is a finite set $R = \{\mathcal{R}_1, \ldots, \mathcal{R}_n\}$ of relations among formulas, called inference rules. For each $\mathcal{R}_i$ there is a unique positive integer j such that $\mathcal{R}_i$ is $(j+1)$–nary relation, i.e. $\mathcal{R}_i$ is a set of $(j+1)$–tuples of formulas. If $(\gamma_1, \ldots, \gamma_j, \sigma) \in \mathcal{R}_i$, then σ is called a direct consequence of $\{\gamma_1, \ldots, \gamma_j\}$ by virtue of $\mathcal{R}_i$. The formulas $\gamma_1, \ldots, \gamma_j$ are called antecedents, σ the consequent.*

Having finite $\Gamma \subset F$ and $\sigma \in F$ we shall say that σ is a consequence of Γ in D (write $\Gamma \vdash \sigma$) iff there exists a sequence of formulas $\alpha_1, \ldots, \alpha_m \in F$ $(m \geq 1)$ such that $\alpha_m = \sigma$ and, for each α_i, either $\alpha_i \in A$ (set of axioms) or $\alpha_i \in \Gamma$ or α_i is a direct consequence of some of the preceding formulas by virtue of some inference rule. Such a sequence is called a deduction of σ from Γ.
We shall say that a deductive system $D = \langle S, F, A, R\rangle$ is regular iff it holds

$$(5.1) \qquad \alpha_1, \ldots, \alpha_{j+1} \in F \quad (\alpha_1, \ldots, \alpha_{j+1}) \in \mathcal{R} \in R \text{ implies } \alpha_i \notin A \text{ for } i = 1, \ldots, j+1\,.$$

In this case it can be easily shown:

$$(5.2) \qquad \sigma \in A \quad \text{iff} \quad \emptyset \vdash \sigma \quad \text{for every} \quad \sigma \in F\,.$$

The presented definition is very general. In [9] the notion of formal theory is concretized by the notion of first–order theory. We give another concrete example here. It is simpler and can serve as a deductive system for describing of ordinary (unconditional) stochastic independence.

Example 1.

(1) We have two classes of symbols: numerals and special symbols. Numerals are usual figures used for notation of nonnegative integers (for example 52 or 0 or 3). Here is the list of seven special symbols: $\emptyset$, ;) (| I

(2) To define the set of formulas we introduce the notion of term. Term is either the symbol $\emptyset$ (empty term) or a finite sequence of distinct numerals separated by commas and ordered in lexicographical ordering i. e. according their numerical values (for example 1,5,11 or 0 or 12,13 but not 3,1). We shall say that terms $\mathcal{A}$ and $\mathcal{B}$ are disjoint iff there is no numeral involved both in $\mathcal{A}$ and $\mathcal{B}$. The formula is a sequence $I(\mathcal{A};\mathcal{B}|\emptyset)$ where $\mathcal{A}$ and $\mathcal{B}$ are disjoint terms.

(3) The axiom is a formula $I(\mathcal{A};\mathcal{B}|\emptyset)$ where either $\mathcal{A}$ or $\mathcal{B}$ is empty.

(4) To describe the inference rules we introduce some operations with terms. Having terms $\mathcal{A}$ and $\mathcal{B}$ we define their conjunction $\mathcal{A}\star\mathcal{B}$ as the term involving exactly those numerals which are involved either in $\mathcal{A}$ or in $\mathcal{B}$ (it is uniquely determined by this requirement). For empty $\mathcal{A}$ and $\mathcal{B}$ the $\mathcal{A}\star\mathcal{B}$ is also empty. We write $\mathcal{B} \subset \mathcal{A}$ if every numeral involved in $\mathcal{B}$ is involved in $\mathcal{A}$. We consider three inference rules. Here are their informal schemes:

$$(\mathcal{R}.1) \qquad I(\mathcal{A};\mathcal{B}|\emptyset) \longrightarrow I(\mathcal{B};\mathcal{A}|\emptyset)$$
$$(\mathcal{R}.2) \qquad I(\mathcal{A};\mathcal{B}|\emptyset) \longrightarrow I(\mathcal{A};\mathcal{C}|\emptyset) \qquad \text{whenever} \quad \mathcal{C} \subset \mathcal{B}$$
$$(\mathcal{R}.3) \qquad I(\mathcal{A};\mathcal{B}|\emptyset)\ I(\mathcal{A}\star\mathcal{B};\mathcal{C}|\emptyset) \longrightarrow I(\mathcal{A};\,\mathcal{B}\star\mathcal{C}|\emptyset)\,.$$

△

A deductive system describes syntactic aspects only. To clarify semantic aspects we need to specify two matters:

- the "field" we want describe by a system under consideration
- the "way" how the theory is related with the field.

Only then we can define the concept of completeness and further important concepts. The "field" will be described by some class of models with distinguished propositions. The aim of a deductive system is to describe formal (logical) relationships among these propositions. The "way" will be realized by interpretations setting up the correspondence between a deductive system and the class of models. Now, we are going to explain our general conception of these matters. Lately, the concrete definitions will be given for our specific fields (for CI definitions 3 and 4, for EMVD definition 5). Our approach is analogical to the classic approach in mathematical logic (see interpretations of first–order theory in [9] chap. 2 sec. 2) but it is slightly modified because we have specific objective.

Definition 2 (syntactic aspects).
Except a deductive system $D = \langle S, F, A, R\rangle$ we consider a class of models $\mathcal{M}$. Every model $M \in \mathcal{M}$ has qualified certain class of propositions $\mathcal{P}_M$ concerning it. Remind that proposition is a statement which has truth value i. e. either TRUE or FALSE.
For every model $M \in \mathcal{M}$ we consider a class of mappings $INT(M)$ called interpretations of M. Every interpretation $\tau \in INT(M)$ has certain set of formulas $D_\tau \subset F$ as its domain and some proposition concerning the model M assigns to every formula from D_τ. We suppose that $\mathcal{P}_M$ is covered by the range τ for every $\tau \in INT(M)$. In case that for every $M \in \mathcal{M}$ and $\tau \in INT(M)$ the range of τ is exactly $\mathcal{P}_M$ we say that the deductive system D is simple.
Given a model $M \in \mathcal{M}$ and its interpretation $\tau \in INT(M)$ we can say that $\sigma \in D_\tau$ is satisfied in M by τ iff the proposition $\tau(\sigma)$ has value TRUE. Having finite $\Gamma \subset F$ and $\sigma \in F$ we shall say that σ is a logical consequence of Γ (write $\Gamma \models \sigma$) iff for every $M \in \mathcal{M}$ and $\tau \in INT(M)$ with $\Gamma \cup \{\sigma\} \subset D_\tau$ it holds:
whenever all formulas Γ are satisfied in M by τ, then also σ is satisfied in M by τ.
A deductive system D is called sound (for a class of models $\mathcal{M}$) iff $\Gamma \vdash \sigma$ entails $\Gamma \models \sigma$ for every finite $\Gamma \subset F$ and $\sigma \in F$. Conversely, it is called complete (for a class of models $\mathcal{M}$) iff $\Gamma \models \sigma$ entails $\Gamma \vdash \sigma$ for every finite $\Gamma \subset F$ and $\sigma \in F$.

Remark 1. For our purposes we can limit ourselves to regular deductive systems only. Indeed, we can replace every original system $D_1 = \langle S, F, A_1, R_1\rangle$ by another regular one $D_2 = \langle S, F, A_2, R_2\rangle$ (with same symbols and formulas) which is equivalent in the following sense:

$$\Gamma \vdash_1 \sigma \quad \text{iff} \quad \Gamma \vdash_2 \sigma \quad \text{whenever finite} \quad \Gamma \subset F,\ \sigma \in F \tag{5.3}$$

To show it we put $A_2 = \{\sigma \in F;\ \emptyset \vdash_1 \sigma\}$. Evidently $A_1 \subset A_2$ and it holds:

$$\alpha_1, \ldots, \alpha_{r+1} \in F \ \ (\alpha_1, \ldots, \alpha_{r+1}) \in \mathcal{R} \in R_1 \quad \alpha_1, \ldots, \alpha_r \in A_2 \text{ implies } \alpha_{r+1} \in A_2\,. \tag{5.4}$$

We assign a finite set of new inference rules $L_{\mathcal{R}}$ to every $\mathcal{R} \in R_1$. Take $\mathcal{R} \subset F^{r+1}$, denote $V = \{1, \ldots, r+1\}$ and put $\mathcal{R}_S = \mathcal{R} \cap K_S$ where $K_S = \prod_{i\in S}\,(F \setminus A_2) \times \prod_{i \in V\setminus S} A_2$ for every $S \subset V$. Clearly, $\{\mathcal{R}_S,\ S \subset V\}$ is a decomposition of $\mathcal{R}$ and by (5.4) $\mathcal{R}_{\{r+1\}} = \emptyset$. Having $S \subset V$ with $s = \operatorname{card} S \geq 1$ define $\mathcal{L}_S \subset (F \setminus A_2)^s$ as the corresponding projection of $\mathcal{R}_S$:

$$\mathcal{L}_S = \Big\{[\sigma_i]_{i\in S}\,;\ \exists\ [\sigma_i]_{i\in V\setminus S}\ \ (\sigma_1, \ldots, \sigma_{r+1}) \in \mathcal{R}_S\Big\}$$

Finally, we put $L_{\mathcal{R}} = \{\mathcal{L}_S;\ S \subset V\ \operatorname{card} S \geq 2\ \&\ (r+1) \in S\}$ and $R_2 = \bigcup\{L_{\mathcal{R}},\ \mathcal{R} \in R_1\}$. Evidently, the new deductive system is regular. The condition (5.3) can be verified with a little effort, too. As D_1 and D_2 have same formulas we can consider the original models and interpretations for D_2. Owing to (5.3) the notions of soudness and completeness of the system are transferred.

Example 2.

Let us consider the deductive system described in Example 1. Every model $M \in \mathcal{M}$ is given by a (finite–valued) random vector $[\xi_i]_{i\in N}$ $(2 \leq \operatorname{card} N < \infty)$. The class $\mathcal{P}_M$ is indexed by couples $\langle A, B\rangle$ of disjoint subsets of N. If both A and B are nonempty then the corresponding proposition $p_M(\langle A, B\rangle)$ express whether $[\xi_i]_{i\in A}$ is stochastically independent of $[\xi_i]_{i\in B}$; $p_M(\langle A, \emptyset\rangle)$ and $p_M(\langle \emptyset, B\rangle)$ have always value TRUE.

Consider a fixed model $M = [\xi_i]_{i\in N}$ and specify the class of its interpretations. Every $\tau \in INT(M)$ is determined by a one–to–one mapping $z : Z \to N$ where Z is a set of numerals with $\operatorname{card} Z = \operatorname{card} N$.

Its domain D_τ is the set of formulas whose terms involve only numerals from Z. As z can be considered as a one–to–one mapping which assigns a subset of N to every such term we define τ as follows:

$$I\,(\mathcal{A};\mathcal{B}|\emptyset) \mapsto^{\tau} p_M\ (\langle z(\mathcal{A}),\, z(\mathcal{B})\rangle)$$

It can be shown using the completeness result from [4] (theorem 3) that the above described deductive system is simple, sound and complete for the described class of models. Note that we can similarly utilize the characterization of fixed–contex CI–statements (see [3], [6]). The reader can find ideas of these procedures in Remark 2. $\triangle$

Now, we are going to deal with the question how to describe CI syntactically. Firstly, we clarify what we shall understand by a deductive system for description CI.

Definition 3 (models for CI).
Models for CI are given by finite–valued random vectors $[\xi_i]_{i\in N}$ where N is an index set. Given a model $M = [\xi_i]_{i\in N}$ we consider the proposition $p_M\,(\langle A,B,C\rangle)$ for every $\langle A,B,C\rangle \in T(N)$. In case that A,B,C are nonempty:
$p_M(\langle A,B,C\rangle)\cdots[\xi_i]_{i\in A}$ is conditionally independent of $[\xi_i]_{i\in B}$ given $[\xi_i]_{i\in C}$.
In case that $C=\emptyset$ and A,B nonempty:
$p_M(\langle A,B,\emptyset\rangle)\cdots[\xi_i]_{i\in A}$ is independent of $[\xi_i]_{i\in B}$.
In case A or B is empty, $p_M(\langle A,B,C\rangle)$ has always value TRUE.
Thus, $\mathcal{P}_M = \{p_M(u);\ u\in T(N)\}$. Every proposition from $\mathcal{P}_M$ will be called CI–statement.

Having two random vectors $[\xi_i]_{i\in K}$ and $[\eta_i]_{i\in L}$ with $\operatorname{card} K \le \operatorname{card} L$ we would like to consider $M = [\xi_i]_{i\in K}$ as a submodel of $\tilde{M} = [\eta_i]_{i\in L}$. Indeed, in this case there exists an injective mapping $v: K\to L$. It can be considered as an injective mapping $\boldsymbol{v}:\ T(K)\to T(L)$

$$\langle A,B,C\rangle \longmapsto^{\boldsymbol{v}} \langle v(A),\, v(B),\, v(C)\rangle \quad \text{whenever} \quad \langle A,B,C\rangle \in T(K)$$

or as an injection $\hat{\boldsymbol{v}}:\ \mathcal{P}_M \to \mathcal{P}_{\tilde{M}}\ \left(p_M(u) \longmapsto^{\hat{\boldsymbol{v}}} p_{\tilde{M}}(\boldsymbol{v}(u)) \text{ for } u\in T(K)\right)$.

Thus, models for CI have some structure. Every reasonable deductive system for describing CI should reflect it in some sense. Now, we are going to motivate and formulate five minimal requirements of such system.

Natural claim is that "interpretations can be transferred to submodels". It means that if certain formulas are interpreted in a model $\tilde{M}$ by some CI–statements then the same formulas can be interpreted as the corresponding CI–statements in a submodel M of $\tilde{M}$ (supposing that M has the corresponding CI–statements). This requirement is exactly formulated by the following condition:

$$(\alpha.1)\quad \begin{cases} \forall\ M=[\xi_i]_{i\in K}\ \ \tilde{M}=[\eta_i]_{i\in L}\ \ \ v: K\to L \text{ injective }\ \tau\in INT(\tilde{M}) \\ \text{there exists } \rho\in INT(M) \text{ such that } \rho_{-1}\left(\mathcal{P}_M\right) = \tau_{-1}\left(\hat{\boldsymbol{v}}(\mathcal{P}_M)\right) \\ \text{and } \hat{\boldsymbol{v}}\circ\rho(\alpha) = \tau(\alpha) \text{ for every } \alpha\in\rho_{-1}\left(\mathcal{P}_M\right). \end{cases}$$

The dual requirement is that "interpretations can be extended to supermodels". Formally:

$$(\alpha.2)\quad \begin{cases} \forall\ M=[\xi_i]_{i\in K}\ \ \tilde{M}=[\eta_i]_{i\in L}\ \ \ v: K\to L \text{ injective }\ \rho\in INT(M) \\ \text{there exists } \tau\in INT(\tilde{M}) \text{ such that } \rho_{-1}\left(\mathcal{P}_M\right) = \tau_{-1}\left(\hat{\boldsymbol{v}}(\mathcal{P}_M)\right) \\ \text{and } \hat{\boldsymbol{v}}\circ\rho(\alpha) = \tau(\alpha) \text{ for every } \alpha\in\rho_{-1}\left(\mathcal{P}_M\right). \end{cases}$$

Further natural claim can be described as follows. If all formulas interpreted as CI–statements by some interpretation ρ can be interpreted by another interpretation τ (perhaps in another model), then

τ can be characterized as an extension of ρ. Formally:

$$(\alpha.3)\quad \left\{ \begin{array}{l} \forall\ M = [\xi_i]_{i\in K} \quad \rho \in INT(M) \quad \tilde{M} = [\eta_i]_{i\in L} \quad \tau \in INT(\tilde{M}) \\ \text{with } \rho_{-1}\left(\mathcal{P}_M\right) \subset \tau_{-1}\left(\mathcal{P}_{\tilde{M}}\right) \text{ there exists } v: K \to L \text{ injective} \\ \text{such that } \rho_{-1}\left(\mathcal{P}_M\right) = \tau_{-1}\left(\hat{\boldsymbol{v}}(\mathcal{P}_M)\right) \text{ and } \hat{\boldsymbol{v}} \circ \rho(\alpha) = \tau(\alpha) \text{ for every } \alpha \in \rho_{-1}\left(\mathcal{P}_M\right). \end{array} \right.$$

Having a set of formulas Γ interpretable as CI–statements in a model $M = [\xi_i]_{i\in N}$ (by $\rho \in INT(M)$) we can naturally introduce its context as the set $\bigcup_{\gamma\in\Gamma} \{[u];\ u \in T(N) \text{ with } p_M(u) = \rho(\gamma)\} \subset N$. Our further requirement is "consistency of interpretations with context". It means that whenever a set of formulas has full context in some model M and can be interpreted by some interpretation τ (in another model $\tilde{M}$), then we can interpret (by τ) formulas describing remaining CI–statements (in $\mathcal{P}_M$). Formally:

$$(\alpha.4)\quad \left\{ \begin{array}{l} \forall\ M = [\xi_i]_{i\in K} \quad \rho \in INT(M) \quad \Gamma \subset F \text{ with } K = \bigcup_{\sigma\in\Gamma} \{[u];\ \rho(\sigma) = p_M(u)\} \\ \forall\ \tilde{M} = [\eta_i]_{i\in L} \quad \tau \in INT(\tilde{M}) \quad \Gamma \subset \tau_{-1}\left(\mathcal{P}_{\tilde{M}}\right) \text{ implies } \rho_{-1}\left(\mathcal{P}_M\right) \subset \tau_{-1}\left(\mathcal{P}_{\tilde{M}}\right). \end{array} \right.$$

The last requirement is "consistency of interpretations with inference rules of the deductive system". It means that is some formula is derivable from interpretable formulas (by some interpretation τ) then it is also interpretable by τ, too. Formally:

$$(\alpha.5)\quad \left\{ \begin{array}{l} \forall\ M = [\xi_i]_{i\in N} \quad \tau \in INT(M) \quad \mathcal{R} \text{ inference rule} \quad \alpha_1, \ldots, \alpha_{r+1} \text{ formulas} \\ (\alpha_1, \ldots, \alpha_{r+1}) \in \mathcal{R},\ \alpha_1, \ldots, \alpha_r \in D_\tau \text{ implies } \alpha_{r+1} \in D_\tau. \end{array} \right.$$

Definition 4 (deductive system for describing CI).
We shall say that a regular deductive system D is a <u>*deductive system for describing CI*</u> *iff for every model M for CI a nonempty collection of mappings $INT(M)$ is given such that*

1. *every $\tau \in INT(M)$ maps certain set D_τ of formulas onto a set of propositions $\mathcal{F}_\tau$ concerning M satisfying $\mathcal{P}_M \subset \mathcal{F}_\tau$*
2. *the conditions $(\alpha.1) - (\alpha.5)$ are satisfied*
3. *D is sound for class of models for CI.*

Moreover, D is <u>*simple*</u> *iff $\mathcal{F}_\tau = \mathcal{P}_M$ for every model M and every $\tau \in INT(M)$.*
D is <u>*complete*</u> *iff it is complete for the class of model for CI.*

The main result of this section says that the existence of a simple and complete deductive system for describing CI is equivalent with the possibility of characterization of CIRs by finite number of regular □–rules. The necessity is contained in the following lemma, the sufficiency in Remark 2.

Lemma 5. *Let there exists a simple and complete deductive system $D = \langle S, F, A, R\rangle$ for describing CI. Then CIRs can be characterized as dependency models closed under a finite set of regular □–rules.*

Proof:

I Firstly we define that set of □–rules.

As D is simple $D_\tau = \tau_{-1}(\mathcal{P}_M)$ for every model M and $\tau \in INT(M)$. Thus, we can correctly define for every index set N:

$\mathcal{K}(N) = \big\{ \boldsymbol{k} : F \to T(N)$ partial mapping ; $\exists\ M = [\xi_i]_{i\in N}\ \ \exists\ \tau \in INT(M)$
such that $D_{\boldsymbol{k}} = D_\tau$ and $\tau(\sigma) = p_M(\boldsymbol{k}(\sigma))$ for every $\sigma \in D_\tau \big\}$.

It makes no problem to see that $\mathcal{K}(N) \neq \emptyset$ and every $\boldsymbol{k} \in \mathcal{K}(N)$ maps $D_{\boldsymbol{k}}$ onto $T(N)$.

Using $(\alpha.1)$ (take $K = L = N$, v identical mapping) we get:

$$\forall M = [\xi_i]_{i\in N} \quad \tau \in INT(M) \text{ iff } \tau = p_M \circ \boldsymbol{k} \text{ for some } \boldsymbol{k} \in \mathcal{K}(N). \tag{5.5}$$

Further, the condition $(\alpha.1) - (\alpha.5)$ can be rewritten:

$$(\beta.1)\quad \begin{cases} \forall\ K,L\ \ v : K \to L \ \text{ injective } \forall\ \boldsymbol{l} \in \mathcal{K}(L)\ \ \exists\ \boldsymbol{k} \in \mathcal{K}(K) \\ \text{such that } D_{\boldsymbol{k}} = \boldsymbol{k}_{-1}\left(\boldsymbol{v}(T(K))\right) \ \text{ and } \boldsymbol{l}(\sigma) = \boldsymbol{v} \circ \boldsymbol{k}(\sigma) \ \text{ for } \sigma \in D_{\boldsymbol{k}} \end{cases}$$

$$(\beta.2)\quad \begin{cases} \forall\ K,L\ \ v : K \to L \ \text{ injective } \ \forall\ \boldsymbol{k} \in \mathcal{K}(K)\ \ \exists\ \boldsymbol{l} \in \mathcal{K}(L) \\ \text{such that } D_{\boldsymbol{k}} = \boldsymbol{l}_{-1}\left(\boldsymbol{v}(T(K))\right) \ \text{ and } \boldsymbol{l}(\sigma) = \boldsymbol{v} \circ \boldsymbol{k}(\sigma) \text{ for } \sigma \in D_{\boldsymbol{k}} \end{cases}$$

$$(\beta.3)\quad \begin{cases} \forall\ \ K,L\ \forall \boldsymbol{k} \in \mathcal{K}(K)\ \ \boldsymbol{l} \in \mathcal{K}(L) \text{ with } D_{\boldsymbol{k}} \subset D_{\boldsymbol{l}}\ \ \exists\ v : K \to L \text{ injective} \\ \text{such that } D_{\boldsymbol{k}} = \boldsymbol{l}_{-1}\left(\boldsymbol{v}(T(K))\right) \text{ and } \boldsymbol{l}(\sigma) = \boldsymbol{v} \circ \boldsymbol{k}(\sigma) \text{ for } \sigma \in D_{\boldsymbol{k}} \end{cases}$$

$$(\beta.4)\quad \begin{cases} \forall\ K,L\ \ \forall \boldsymbol{k} \in \mathcal{K}(K) \text{ and } \Gamma \subset F \text{ with } K = \bigcup_{\sigma\in\Gamma} [\boldsymbol{k}(\sigma)] \\ \text{whenever } \boldsymbol{l} \in \mathcal{K}(L) \text{ satisfies } \Gamma \subset D_{\boldsymbol{l}}, \ \text{ then } D_{\boldsymbol{k}} \subset D_{\boldsymbol{l}} \end{cases}$$

$$(\beta.5)\quad \begin{cases} \forall\ N\ \ \forall\ \boldsymbol{k} \in K(N)\ \ \forall\ \mathcal{R} \in R \quad \alpha_1, \ldots, \alpha_{r+1} \in F\ \ (r \geq 1) \\ (\alpha_1, \ldots \alpha_{r+1}) \in \mathcal{R} \text{ and } \alpha_1, \ldots, \alpha_r \in D_{\boldsymbol{k}} \text{ implies } \alpha_{r+1} \in D_{\boldsymbol{k}}\,. \end{cases}$$

The first $\square$–rule has no antecedents ($r = 0$) and corresponds to the set of axioms A:
N index set $\ \longrightarrow\ \boldsymbol{A}_0(N) \equiv \bigcup_{\boldsymbol{k}\in\mathcal{K}(N)} \boldsymbol{k}\left(A \cap D_{\boldsymbol{k}}\right) \subset T(N)$

Every other $\square$–rule corresponds to some inference rule $\mathcal{R}_j \subset F^{r+1}$ (here $r \geq 1$):
N index set $\ \longrightarrow\ \boldsymbol{A}_j(N) \equiv \bigcup_{\boldsymbol{k}\in\mathcal{K}(N)} \overline{\boldsymbol{k}}\left(\mathcal{R}_j \cap D_{\boldsymbol{k}}^{r+1}\right) \subset T(N)^{r+1}$
where $\overline{\boldsymbol{k}} :\ D_{\boldsymbol{k}}^{r+1} \longrightarrow T(N)^{r+1}$ is defined by $\overline{\boldsymbol{k}}\,(\sigma_1, \ldots, \sigma_{r+1}) = (\boldsymbol{k}(\sigma_1), \ldots, \boldsymbol{k}(\sigma_{r+1}))$

It makes no problem to verify using $(\beta.1)$, $(\beta.2)$ that the above defined $\square$–rules are regular. $\otimes$

II Let K be an index set, $\boldsymbol{k} \in \mathcal{K}(K)$, $\Gamma \subset F$ with $K = \bigcup_{\sigma\in\Gamma} [\boldsymbol{k}(\sigma)]$. Then for each model $M = [\eta_i]_{i\in L}$ and $\tau \in INT(M)$ with $\Gamma \subset D_\tau$ there exists $v : K \to L$ injective such that $\tau(\sigma) = p_M(\boldsymbol{v}(\boldsymbol{k}(\sigma)))$ whenever $\sigma \in D_{\boldsymbol{k}}$.

Indeed: Apply (5.5) to $M = [\xi_i]_{i\in L}$ and $\tau \in INT(M)$ and find $\boldsymbol{l} \in \mathcal{K}(L)$ with $\tau = p_M \circ \boldsymbol{l}$. As $\Gamma \subset D_{\boldsymbol{l}}$ we can use $(\beta.4)$ to derive $D_{\boldsymbol{k}} \subset D_{\boldsymbol{l}}$. Then use $(\beta.3)$. $\otimes$

III Having fixed index set N it holds $\boldsymbol{A}_0(N) = T(N) \setminus T_*(N)$.

Indeed: Let $u \in \boldsymbol{A}_0(N)$ i.e. $\exists\ \boldsymbol{k} \in \mathcal{K}(N)\ \ \exists\ \sigma \in A \cap D_{\boldsymbol{k}}\ \ \boldsymbol{k}(\sigma) = u$. By Lemma 4 (for $V = \emptyset$) and Lemma 1 we find a model $M = [\xi_i]_{i\in N}$ such that $p_M(w)$ has value TRUE iff $w \in T(N) \setminus T_*(N)$. Further, $\tau \equiv p_M \circ \boldsymbol{k} \in INT(M)$ by (5.5) and $\sigma \in D_\tau$. But $\sigma \in A$ implies $\emptyset \vdash \sigma$ and it gives $\emptyset \models \sigma$ (D is sound). The definition of $\emptyset \models \sigma$ (see Def. 2) applied to M and τ says $p_M(u) \to$ TRUE. Hence, $u \in T(N) \setminus T_*(N)$.

Conversely, let $u \in T(N) \setminus T_*(N)$; put $K = [u]$. Fix some $\boldsymbol{l} \in \mathcal{K}(N)$ and choose $\sigma \in D_{\boldsymbol{l}}$ with $\boldsymbol{l}(\sigma) = u$. Take identical mapping $i : K \to N$ and by $(\beta.1)$ find $\boldsymbol{k} \in \mathcal{K}(K)$ with $\sigma \in D_{\boldsymbol{k}}$ and $\boldsymbol{k}(\sigma) = \boldsymbol{i}_{-1}(u)$. Consider any model $M = [\eta_i]_{i\in L}$ and $\tau \in INT(M)$ with $\sigma \in D_\tau$. By II ($\Gamma = \{\sigma\}$) there exists $v :\ K \to L$ injective such that $\tau = p_M \circ \boldsymbol{v} \circ \boldsymbol{i}_{-1}$. As u is trivial, $\boldsymbol{v} \circ \boldsymbol{i}_{-1}(u)$ is trivial, too. Hence $\tau(\sigma) \to$ TRUE (see Def. 3). Thus, we have verified $\emptyset \models \sigma$. It gives $\emptyset \vdash \sigma$ (D is complete) and consecutively $\sigma \in A$ (D is regular see (5.2)). Hence $u \in \bigcup_{\boldsymbol{l}\in\mathcal{K}(N)} \boldsymbol{l}(A \cap D_{\boldsymbol{l}}) = \boldsymbol{A}_0(N)$. $\otimes$

IV Given index set N every $I \in CIR(N)$ is closed under all defined $\square$–rules.

Indeed: Every CIR contains $T(N) \setminus T_*(N)$. Hence, by III I is closed under the first $\square$–rule $\boldsymbol{A}_0(N)$. Consider a $\square$–rule given by $\boldsymbol{A}_j(N) \subset T(N)^{r+1}$ ($r \geq 1$ corresponding $\mathcal{R}_j \in R$). Suppose $u_1, \ldots, u_{r+1} \in T(N)$ with $u_1, \ldots, u_r \in I$ and $(u_1, \ldots, u_{r+1}) \in \boldsymbol{A}_j(N)$. By the definition of $\boldsymbol{A}_j(N)$ there exists $\boldsymbol{k} \in \mathcal{K}(N)$ and $(\sigma_1, \ldots, \sigma_{r+1}) \in \mathcal{R}_j$ with $\sigma_i \in D_{\boldsymbol{k}}$, $\boldsymbol{k}(\sigma_i) = u_i$ $(i = 1, \ldots, r+1)$. Since $I \in CIR(N)$ there exists $M = [\xi_i]_{i \in N}$ such that

$$p_M(w) \longrightarrow \text{ TRUE} \quad \text{iff } w \in I \qquad \text{whenever } w \in T(N). \tag{5.6}$$

Take $M = [\xi_i]_{i \in N}$ and $\tau = p_M \circ \boldsymbol{k} \in INT(M)$ (see (5.5)). As $u_1, \ldots, u_r \in I$ we have $p_M(u_i) \to$ TRUE for $i = 1, \ldots, r$. But $\{\sigma_1, \ldots, \sigma_r\} \vdash \sigma_{r+1}$ implies $\{\sigma_1, \ldots, \sigma_r\} \models \sigma_{r+1}$ (D is sound). Apply it to M and τ and get $p_M(u_{r+1}) \to$ TRUE i.e. $u_{r+1} \in I$ by (5.6). $\otimes$

V Let N be an index set and $I \subset T(N)$ is closed under all defined $\square$–rules. Then $I \in CIR(N)$.

Indeed: By Lemma 1 it suffices to find (for every $u \in T(N) \setminus I$) some model $\tilde{M} = [\eta_i]_{i \in N}$ such that:

$$p_{\tilde{M}}(u) \longrightarrow \text{ FALSE} \quad \text{and } p_{\tilde{M}}(w) \longrightarrow \text{ TRUE} \quad \text{for every } w \in I. \tag{5.7}$$

Take $\boldsymbol{k} \in \mathcal{K}(N)$, choose $\Gamma \subset D_{\boldsymbol{k}}$ finite with $\boldsymbol{k}(\Gamma) = I$ and $\sigma \in D_{\boldsymbol{k}}$ with $\boldsymbol{k}(\sigma) = u$. We shall verify that:

$$\Gamma \not\vdash \sigma. \tag{5.8}$$

By contradiction: Let $\alpha_1, \ldots, \alpha_m$ be a deduction of σ from Γ. Put $\tilde{\Gamma} = \boldsymbol{k}_{-1}(I)$. Find minimal α_s which does not belong to $\tilde{\Gamma} \cup A$ (note that $\sigma \notin A$, otherwise the definition of $\boldsymbol{A}_0(N)$ gives $u \in \boldsymbol{A}_0(N) \subset I$). Clearly, α_s is a direct consequence of preceding formulas by virtue some $\mathcal{R}_j \in R$. As D is regular, they belong to $\tilde{\Gamma}$. Thus, we found $\beta_1, \ldots, \beta_r \in \tilde{\Gamma}$ and $\beta_{r+1} \in F \setminus \tilde{\Gamma}$ such that $(\beta_1, \ldots, \beta_{r+1}) \in \mathcal{R}_j$. As $\beta_1, \ldots, \beta_r \in \tilde{\Gamma} \subset D_{\boldsymbol{k}}$ by (β.5) we derive $\beta_{r+1} \in D_{\boldsymbol{k}}$. Hence $(\boldsymbol{k}(\beta_1), \ldots, \boldsymbol{k}(\beta_{r+1})) \in \boldsymbol{A}_j(N)$. As $\boldsymbol{k}(\beta_1), \ldots, \boldsymbol{k}(\beta_r) \in I$ and I is closed under all $\square$–rules we derive $\boldsymbol{k}(\beta_{r+1}) \in I$ i.e. $\beta_{r+1} \in \tilde{\Gamma}$ – it contradicts the definition β_{r+1}.

But (5.8) implies $\Gamma \not\models \sigma$ (D is complete). By the definition of $\Gamma \not\models \sigma$ we find $M = [\xi_i]_{i \in L}$ and $\tau \in INT(M)$ with $\Gamma \cup \{\sigma\} \subset D_\tau$ such that $\tau(\sigma) \to$ FALSE and $\tau(\gamma) \to$ TRUE for $\gamma \in \Gamma$. By III $T(N) \setminus T_*(N) = \boldsymbol{A}_0(N) \subset I$ and hence $N = \bigcup_{u \in I}[u] = \bigcup_{\sigma \in \Gamma}[\boldsymbol{k}(\sigma)]$. We can use II (for $K = N$) to find $v : N \to L$ injective such that $\tau(\gamma) = p_M \circ \boldsymbol{v} \circ \boldsymbol{k}(\gamma)$ for $\gamma \in D_{\boldsymbol{k}}$. Consider a new model $\tilde{M} = \left[\xi_{v(i)}\right]_{i \in N}$. Evidently, it holds

$$p_{\tilde{M}}(w) \longrightarrow \text{ TRUE} \quad \text{iff} \quad p_M(\boldsymbol{v}(w)) \longrightarrow \text{ TRUE} \qquad \text{for every } w \in T(N). \tag{5.9}$$

For each $w \in I$ we take $\gamma \in \Gamma$ with $w = \boldsymbol{k}(\gamma)$. As $p_M \circ \boldsymbol{v} \circ \boldsymbol{k}(\gamma) = p_M \circ \boldsymbol{l}(\gamma) = \tau(\gamma) \to$ TRUE by (5.9) we get $p_{\tilde{M}}(w) = p_{\tilde{M}} \circ \boldsymbol{k}(\gamma) \to$ TRUE. Analogously, $p_{\tilde{M}}(u) \to$ FALSE. Thus, $\tilde{M}$ satisfies (5.7). $\triangle$

Consequence 2. *There is no simple and complete deductive system for describing CI.*

Proof. Combine Consequence 1 and Lemma 5. $\triangle$

Remark 2. In Example 2 we mentioned that a characterization for some subclass of CI–statements can be utilize for its syntactic description. The following consideration underlines the principles of these procedures. It shows how a hypothetic simple syntactical description of CI would have looked.

Let us suppose that we have a characterization of CIRs by means of finite number of regular $\square$–rules $(\mathcal{A}.0) - (\mathcal{A}.s)$ where

a) every $(\mathcal{A}.r)$ is given by $\boldsymbol{A}_r(N) \subset T(N)^{r+1}$ $(r = 0, \ldots, s \quad N$ index set)

b) $\boldsymbol{A}_0(N) = T(N) \setminus T_*(N)$

c) $\boldsymbol{A}_r(N) \subset T_*(N)^{r+1}$ whenever $r \geq 1$

d) $r \geq 1,\ (u_1, \ldots, u_{r+1}) \in A_r(N)$ implies $[u_{r+1}] \subset \bigcup_{i=1}^{r} [u_i]$.

These additional demands would be partially justified by Proposition 2 (although we have not proved that the existence of a finite characterization implies the existence of the special one). Now we can construct a simple syntactic description of CI as follows.

Deductive system

The set of symbols is the same as in Example 1. We undertake the notion of term also. Consider formulas of the form $I(\mathcal{A};\, \mathcal{B}|\mathcal{C})$ where $\mathcal{A}, \mathcal{B}, \mathcal{C}$ are pairwise disjoint terms. To specify axioms and inference rules we introduce (for every N) the class $\kappa(N)$ as the collection of all one–to–one mappings $z : Z \to N$ where Z is a set of numerals with $\operatorname{card} Z = \operatorname{card} N$. Note that every element $z \in \kappa(N)$ can be considered as a partial mapping $\boldsymbol{z} : F \to T(N)$ (the domain $D_{\boldsymbol{z}}$ is the set of formulas involving variables from $z_{-1}(N)$ and $I(\mathcal{A};\, \mathcal{B}|\mathcal{C}) \mapsto^{\boldsymbol{z}} \langle z(\mathcal{A}),\, z(\mathcal{B}),\, z(\mathcal{C})\rangle$), i.e. an element of $\mathcal{K}(N)$ – see the proof of Lemma 5. Moreover, it can be considered as a mapping $\overline{\boldsymbol{z}} :\ F^{r+1} \to T(N)^{r+1}$ for $r \geq 1 :\ \overline{\boldsymbol{z}}\,(\alpha_1, \ldots, \alpha_{r+1}) = (\boldsymbol{z}(\alpha_1), \ldots, \boldsymbol{z}(\alpha_{r+1}))$. Then we specify the set of axioms:

$$A = \bigcup_{N} \bigcup_{z \in \kappa(N)} (\boldsymbol{z}_{-1})\, \boldsymbol{A}_0(N) \subset F$$

Further, for every $r = 1, \ldots, s$ we define an inference rule

$$\mathcal{R}_r = \bigcup_{N} \bigcup_{z \in \kappa(N)} (\overline{\boldsymbol{z}}_{-1}) \boldsymbol{A}_r(N) \subset F^{r+1}$$

Models and interpretations

The class of models is described in Def. 3. Given a model $M = [\xi_i]_{i \in N}$ every its interpretation $\tau \in INT(M)$ is given by an element $z \in \kappa(N)$:
$D_\tau = D_{\boldsymbol{z}}$ and $\tau(\sigma) = p_M(\boldsymbol{z}(\sigma))$ for $\sigma \in D_\tau$, i.e. $I(\mathcal{A};\, \mathcal{B}|\mathcal{C}) \mapsto^{\tau} p_M\,(\langle z(\mathcal{A}),\, z(\mathcal{B}),\, z(\mathcal{C})\rangle)$.

I The deductive system is regular.
Indeed: It easily follows from b):
$I(\mathcal{A};\, \mathcal{B}|\mathcal{C}) \in A \Longleftrightarrow \mathcal{A}$ or $\mathcal{B}$ is empty term.
Hence, c) gives: $(\sigma_1, \ldots, \sigma_{r+1}) \in \mathcal{R}_r,\ r \geq 1 \Longrightarrow \sigma_i \notin A$ for $i = 1, \ldots, r+1$. $\otimes$

II $(\alpha.1) - (\alpha.5)$ holds.
Indeed: Clearly, the conditions $(\beta.1) - (\beta.5)$ mentioned in the proof of Lemma 5 are satisfied ($(\beta.5)$ follows from d)). Hence, $(\alpha.1) - (\alpha.5)$ can be easily derived. $\otimes$

III Let $(\gamma_1, \ldots, \gamma_{r+1}) \in \mathcal{R}_r \quad M = [\xi_i]_{i \in N} \quad \tau \in INT(M)$ with $\gamma_1, \ldots, \gamma_{r+1} \in D_\tau$
$\tau(\gamma_1), \ldots, \tau(\gamma_r) \to$ TRUE. Then $\tau\,(\gamma_{r+1}) \to$ TRUE.
Indeed: Consider $k \in \kappa(N)$ given by the equality $\tau = p_M \circ \boldsymbol{k}$.
Put $I = \{u \in T(N);\, p_M(u) \to \text{TRUE}\}$. Evidently $\boldsymbol{k}(\gamma_1), \ldots, \boldsymbol{k}(\gamma_r) \in I$. As $I \in CIR(N)$ it is closed under $(\mathcal{A}.r)$. Hence, to derive $\tau\,(\gamma_{r+1}) \to$ TRUE i.e. $\boldsymbol{k}\,(\gamma_{r+1}) \in I$ we need to show $(\boldsymbol{k}(\gamma_1), \ldots, \boldsymbol{k}(\gamma_{r+1})) \in \boldsymbol{A}_r(N)$. As $(\gamma_1, \ldots, \gamma_{r+1}) \in \mathcal{R}_r$ there exists an index set $L,\ z \in \kappa(L)$ and $(w_1, \ldots, w_{r+1}) \in \boldsymbol{A}_r(L)$ such that $\boldsymbol{z}(w_i) = \gamma_i$. We can suppose $L = \bigcup_{i=1}^{r+1} [w_i]$ (by $(\beta.1)$ and regularity of $(\mathcal{A}.r)$). Further, using $(\beta.4)$ and $(\beta.3)$ we find $v : L \to N$ injective such that $\boldsymbol{k}(\gamma) = \boldsymbol{v} \circ \boldsymbol{z}(\gamma)$ for $\gamma \in D_{\boldsymbol{k}}$. Hence, by regularity of $(\mathcal{A}.r)$ $(\boldsymbol{k}(\gamma_1), \ldots, \boldsymbol{k}(\gamma_{r+1})) \in \boldsymbol{A}_r(N)$. $\otimes$

IV $\Gamma \vdash \sigma$ implies $\Gamma \models \sigma$ $(\Gamma \subset F$ finite, $\sigma \in \Gamma)$

Indeed: It can be derived from $(\alpha.5)$ and regularity of the deductive system:

$$(5.10) \qquad \Gamma \vdash \sigma \Longrightarrow \forall\, M \text{ model } \forall\, \tau \in INT(M) \text{ with } \Gamma \subset D_\tau \qquad \sigma \in D_\tau \cup A$$

Consider $M = [\xi_i]_{i\in N}$ and $\tau \in INT(M)$ with $\Gamma \cup \{\sigma\} \subset D_\tau$ and $\tau(\sigma) \to$ TRUE for $\gamma \in \Gamma$. It suffices to show $\tau(\sigma) \to$ TRUE. Thus, take a deduction $\alpha_1, \ldots, \alpha_m$ of σ from Γ and prove by induction that $\tau(\alpha_i) \to$ TRUE $i = 1, \ldots, m$. If $\alpha_i \in \Gamma$ it is evident. If $\alpha_i \in A$, then $\tau(\alpha_i) = p_M(u)$ for a trivial triplet u and $p_M(u) \to$ TRUE by definition. Let α_i be a direct consequence of preceding formulas by virtue of some inference rule $\mathcal{R}_r$. They are not axioms by I and by (5.10) belong to D_τ. We can use III to derive $\tau(\alpha_i) \to$ TRUE. $\otimes$

V $\Gamma \models \sigma$ implies $\Gamma \vdash \sigma$ $(\Gamma \subset F$ finite, $\sigma \in \Gamma)$

Indeed: As Γ is finite we easily find an index set N and $z \in \kappa(N)$ such that $\Gamma \cup \{\sigma\} \subset D_{\boldsymbol{z}}$. It can be easily derived from $\Gamma \models \sigma$ and the definition of interpretations:

$$(5.11) \qquad \forall\, I \in CIR(N) \qquad \boldsymbol{z}(\Gamma) \subset I \qquad \text{implies} \qquad \boldsymbol{z}(\sigma) \in I$$

Further, we define

$K = \{u \in T(N);$ there exists a sequence $u_1, \ldots, u_m \in T(N)$ where $u_m = u$ and for each u_i either $u_i \in \boldsymbol{z}(\Gamma)$ or $(u_{i_1}, \ldots, u_{i_r}, u_i) \in \boldsymbol{A}_r(N)$ for some $r \geq 0$ and some preceding triplets$\}$

Evidently K is closed under $(\mathcal{A}.0) - (\mathcal{A}.s)$ and contains $\boldsymbol{z}(\Gamma)$. Hence $K \in CIR(N)$ and by (5.11) contains $\boldsymbol{z}(\sigma)$. Let $u_1, \ldots, u_m$ be the corresponding sequence. Put $\alpha_i = \boldsymbol{z}_{-1}(u_i)$ $i = 1, \ldots, m$. It makes no problem to verify (using the definitions of A and $\mathcal{R}_r$) that $\alpha_1, \ldots, \alpha_m$ is a deduction of σ from Γ.$\otimes$

VI We can summarize I - V: the constructed deductive system would give simple syntactic description of CI. $\triangle$

6 ANALOGY WITH EMVD

As we have mentioned that questions concerning CI–statements have analogy in the theory of relational databases namely in questions concerning *embedded multivalued dependencies* (EMVDs). This model is specified by the following definition.

Definition 5 (models for EMVD).
Every model is given by an index set N $(2 \leq \mathrm{card}\, N < \infty)$, by a collection of nonempty sets $\{X_i;\ i \in N\}$ and by a nonempty subset $\mathsf{R} \subset \prod_{i\in N} X_i$ (R is called database relation).
Given $\langle A, B, C\rangle \in T(N)$ we write $C \twoheadrightarrow A \,|\, B$ in R iff

$$(6.1) \qquad \forall\, x, y \in \mathsf{R} \text{ with } x_C = y_C \quad \exists\, z \in \mathsf{R} \qquad z_{A\cup C} = x_{A\cup C} \;\&\; z_{B\cup C} = y_{B\cup C}$$

(here x_C denotes $[x_i]_{i\in C}$ whenever $C \subset N$ and $x = [x_i]_{i\in N} \in \prod_{i\in N} X_i$).
We speak about EMVD–statements.

Remark. EMVD–statements are usually defined (see [15]) for triplets $\langle A, B, C\rangle$ satisfying $A \cap B \subset C$ (by the same requirement). But it holds: $C \twoheadrightarrow A \,|\, B$ in R if and only if

$C \twoheadrightarrow (A \setminus C) \,|\, (B \setminus C)$ in R for every such triplet and database relation R. Thus, our definition is not restrictive.

A lot of effort was exerted to characterize formal relationships among EMVD–statements. For example in [1] multivalued dependencies (subclass of EMVDs $C \twoheadrightarrow A \,|\, B$ with $N = A \cup B \cup C$) were axiomatized. Analogical result for crosses (EMVDs $C \twoheadrightarrow A \,|\, B$ with $C = \emptyset$) is in [11]. The article [15] gives a method how test inferring among Z–EMVDs (EMVDs $C \twoheadrightarrow A \,|\, B$ with fixed B) and shows that the class of all EMVDs has no complete axiomatization. Note that although all these authors speak about axiomatization they did not give a formal axiomatic theory in sense of mathematical logic. By axiomatization they understand a characterization by means of a finite number of $\Box$–rules (in our terminology).

Considering a natural correspondence $p_M(\langle A, B, C\rangle) \cdots C \twoheadrightarrow A|B$ above mentioned results very resemble the results concerning special subclasses of CI–statements ([3], [4], [6], [7]):
It leads to a hypothesis that formal relationships among CI–statements and those among EMVD–statements are identical. More precisely, the hypothesis can be formulated as follows:

$$(?) \quad \left\{ \begin{array}{l} \text{We conjecture that for every index set } N \text{ and } I \subset T(N) \\ I \in CIR(N) \;\text{ iff }\; I = \{\langle A, B, C\rangle \in T(N);\, \twoheadrightarrow A \,|\, B \text{ in } \mathsf{R}\} \;\text{ for some} \\ \text{database relation } \mathsf{R} \text{ on } N \; \big(\text{i.e. } \mathsf{R} \subset \prod_{i \in N} X_i\big)\,. \end{array} \right.$$

Many arguments support this hypothesis:

- it is confirmed in case card $N \leq 3$
- in some special case the mentioned concepts coincide. Indeed, let $M = [\xi_i]_{i \in N}$ be a random vector (ξ_i takes values in X_i) such that every marginal of its distribution P is uniformly distributed on its support. Define $\mathsf{R} \subset \prod_{i \in N} X_i$ as the support of P :
  $\mathsf{R} = \big\{x \in \prod_{i \in N} X_i;\; P(x) > 0\big\}$. Then it holds:
  $p_M\,(\langle A, B, C\rangle) \longrightarrow \text{ TRUE } \;\text{ iff }\; C \twoheadrightarrow A|B \text{ in } \mathsf{R} \quad \text{whenever } \langle A, B, C\rangle \in T(N)$
- the characterization for MVDs (see [1]) is identical with the characterization of the corresponding class of CI–statements i. e. fixed–contex ([3], [6]).
- analogical case occurs for crosses (see [11]) and marginal CI–statements ([4], [7])
- even our result (Consequence 1) is analogical to negative result from [15]. We have derived (in Proposition 1) the same property for CI–statements as Sagiv and Walecka for EMVD–statements.

Nevertheless, we refuse both implications in our hypothesis (?). Example 3 disclaims sufficiency, Example 4 necessity.

Example 3. ($\Box$–rule holds for $CI \not\Rightarrow$ it holds for EMVD)
It was shown in [16] that every CIR is closed under the following $\Box$–rule:

$$(6.2) \qquad [I(A;\, B \,|\, C \cup D) \;\&\; I(C;\, D \,|\, A) \;\&\; I(C, D \,|\, B) \;\&\; I(A;\, B \,|\, \emptyset)] \;\longrightarrow\; I(C;\, D \,|\, \emptyset).$$

But is fails in the case of EMVD. Take $N = \{1, 2, 3, 4\}$, $X_i = \{0, 1\}$, for $i \in N$, and $\mathsf{R} \subset \prod_{i \in N} X_i$ given by the following list:

$$\begin{array}{c} (\,0, \;\; 0, \;\; 0 \;\; 0\,) \\ (\,0, \;\; 0, \;\; 0 \;\; 1\,) \\ (\,0, \;\; 1, \;\; 0 \;\; 0\,) \\ (\,1, \;\; 0, \;\; 0 \;\; 0\,) \\ (\,1, \;\; 1, \;\; 0 \;\; 0\,) \\ (\,1, \;\; 1, \;\; 1 \;\; 0\,) \end{array}$$

It makes no problem to verify that EMVDs $\{3,4\} \twoheadrightarrow \{1\} \mid \{2\}$, $\{1\} \twoheadrightarrow \{3\} \mid \{4\}$, $\{2\} \twoheadrightarrow \{3\} \mid \{4\}$, $\emptyset \twoheadrightarrow \{1\} \mid \{2\}$ hold in R. But EMVD $\emptyset \twoheadrightarrow \{3\} \mid \{4\}$ does not hold in R. Thus, $I = \{\langle A, B, C\rangle ; C \twoheadrightarrow A \mid B \text{ in } \mathsf{R}\}$ is a dependency model given by a database relation which is not CIR. △

Example 4. (□–rule holds for EMVD $\not\Rightarrow$ it holds for CI)
Firstly, we shall show that every dependency I model given by a database relation satisfies the following □–rule:

(6.3) $$[I(A;\, B \mid C \cup D) \;\&\; I(C;\, D \mid A) \;\&\; I(C;\, D \mid B)] \;\longrightarrow\; I(C;\, D \mid A \cup B)$$

Indeed: Consider $\mathsf{R} \subset \prod_{i \in N} X_i$ satisfying antecedents of (6.3). We want verify $A \cup B \twoheadrightarrow C \mid D$ in R. Thus, take $x, y \in \mathsf{R}$ with $x_{A\cup B} = y_{A\cup B}$ and construct $z \in \mathsf{R}$ with $z_{A\cup B\cup C} = x_{A\cup B\cup C}$ and $z_{A\cup B\cup D} = y_{A\cup B\cup D}$. As $x_A = y_A$ by $A \twoheadrightarrow C \mid D$ we find $v \in \mathsf{R}$ with $v_{A\cup C\cup D} = [x_A,\, x_C,\, y_D] = [y_A,\, x_C,\, y_D]$. As $x_B = y_B$ by $B \twoheadrightarrow C \mid D$ we find $w \in \mathsf{R}$ with $w_{B\cup C\cup D} = [x_B,\, x_C,\, y_D] = [y_B,\, x_C,\, y_D]$. As $v_{C\cup D} = w_{C\cup D}$ by $C \cup D \twoheadrightarrow A \mid B$ we find $z \in \mathsf{R}$ with $z_{A\cup C\cup D} = v_{A\cup C\cup D}$ and $z_{B\cup C\cup D} = w_{B\cup C\cup D}$. Evidently $z_{A\cup B\cup C\cup D} = [x_A,\, x_B,\, x_C,\, y_D] = [y_A,\, y_B,\, x_C,\, y_D]$. ⊗

But (6.3) fails in case CI. Take $N = \{1,2,3,4\}$, $X_i = \{0,1\}$ for $i \in N$ and define a probability measure on $\prod_{i \in N} X_i$ as follows:

$$\begin{array}{lcl}
(0,\; 0,\; 0\;\; 0) & \longrightarrow & 0.2 \\
(0,\; 0,\; 1\;\; 0) & \longrightarrow & 0.2 \\
(0,\; 0,\; 0\;\; 1) & \longrightarrow & 0.2 \\
(0,\; 0,\; 1\;\; 1) & \longrightarrow & 0.1 \\
(0,\; 1,\; 1\;\; 1) & \longrightarrow & 0.1 \\
(1,\; 0,\; 1\;\; 1) & \longrightarrow & 0.1 \\
(1,\; 1,\; 1\;\; 1) & \longrightarrow & 0.1
\end{array}$$

The corresponding CIR contains triplets $\langle\{1\},\, \{2\},\, \{34\}\rangle$, $\langle\{3\},\, \{4\},\, \{1\}\rangle$ and $\langle\{3\},\, \{4\},\, \{2\}\rangle$ but it does not contain $\langle\{3\},\, \{4\},\, \{12\}\rangle$. By (6.3) it is not a dependency model given by a database relation. △

7 CONCLUDING REMARKS

All this paper was more or less engaged in the following problem:

Can we describe formal properties of CIRs by means of a complete formal axiomatic theory?

In the fifth section we have shown that syntactic description by means of a simple deductive system (i. e. a system whose formulas correspond to individual CI–statements) is equivalent to characterization of CIRs by means of finite number of □–rules.

Existence result in Proposition 2 speaks about characterization of CIRs by means of countably many □–rules. But it can be utilized only in case when one would have had a “list” of all CIRs at his disposal.

By Consequence 1 CIRs cannot be characterized by finite number of □–rules and hence they cannot be described by a simple deductive system, especially by such a formal theory. Nevertheless our result does not refuse the possibility to describe formal properties of CIRs by a complete deductive system which is not simple. If we forsake this demand and allow that the system can have a wider class of formulas (for example formulas can correspond with finite conjunctions of CI–statements) then our arguments cannot be used since all our new formal properties can be embraced in one inference rule! Then our result (and similarly the analogical result [15] concerning EMVDs) gives only relatively negative conclusion: CIRs cannot be characterized in classic way.

Thus, the plan of further investigation is to seek for a more wider formal system describing CIRs. Some attempt in this respect was made in [17] (the concept of M–relation).

Examples 3 and 4 document that our problem is not equivalent to an analogical problem in the theory of relational databases namely decidability of testing implications of EMVDs (see [15]).

Acknowledgement:
I am grateful to my colleagues F. Matúš and I. Kramosil for their valuable comments.

REFERENCES

[1] Beeri C., Fagin R. and Howard J. H.: A complete axiomatization for functional and multivalued dependencies in database relations. Proc. ACM SIGMOD Int. Conf. on Management of Data, Toronto 1977, pp. 47–61.

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

[3] Geiger D. and Pearl J.: Logical and algorithmic properties of conditional independence. Technical report CSD 870056, R–97–II, University of California, Los Angeles, September 1988; also in Proc. 2nd International Workshop on AI and Statistics, January 4–7, 1989, Fort Lauderdale, Florida.

[4] Geiger D., Paz A. and Pearl J.: Axioms and algorithms for inference involving probabilistic independence. Technical report CSD 890031, R–119–I, University of California, Los Angeles, October 1989; also to appear in Information and Computation 1990.

[5] Kramosil I.: A note on nonaxiomatizability of independence relations generated by certain probabilistic structure. Kybernetika vol. 24 (1988), n. 6, pp. 439–446.

[6] Malvestuto F. M.: A unique formal system for binary decompositions of database relations, probability distributions and graphs. To appear in Information Science 1992.

[7] Matúš F.: Independence and Radon projection on compact groups. (in Slovak), PhD. Thesis, Prague 1988.

[8] Matúš F.: Ascending and descending conditional independence relations. in Transactions of the 11–th Prague Conference on Information Theory, Statistical Decision Functions and Random Processes vol.B., Academia, Prague 1992, pp. 189–200.

[9] Mendelson E.: Introduction to Mathematical Logic, 2nd edition, D. van Nostrand Comp. New York Cincinnati Toronto London Melbourne 1979.

[10] Mouchart M. and Rolin J.–M.: A note on conditional independence with statistical applications. Statistica vol. XLIV (1984), n. 4, pp. 557–584.

[11] Paredaens J.: The interaction of integrity constrains in an information system. Journal of Computer and System Sciences vol. 20 (1980), n. 3, pp. 310–329.

[12] Pearl J.: Markov and Bayes networks: a comparison of two graphical representations of probabilistic knowledge. Technical report CSD 860024, R–46–I, University of California, Los Angeles, October 1986.

[13] Perez A. and Jiroušek R.: Constructing and intensional expert system (INES). Medical Decision Making: diagnostic strategies and expert systems, North Holland 1985, pp. 307–315.

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

[15] Sagiv Y. and Walecka S. F.: Subset dependencies and a completeness result for a subclass of embedded multivalued dependencies. Journal of the Association for Computing Machinery vol. 29 (1982), n. 1, pp. 103–117.

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

[17] Studený M.: Attempts at axiomatic description of conditional independence. Proc. Workshop on Uncertainty Processing in Expert Systems, June 20–23, 1988, Alšovice, Czechoslovakia; also in Kybernetika vol. 25 (1989), supplement to n. 3, pp. 72–79.

[18] Studený M.: Multiinformation and conditional independence I. Technical report n. 1619, Institute of Information Theory and Automation, Prague, October 1989.

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

tel 42-2-66052304
fax 42-2-66414903
email studeny@cspgas11.bitnet