Characteristic Properties of Equivalent Structures in Compositional Models

Václav Kratochvíl

Institute of Information Theory and Automation, Acad. Sci. of the Czech Republic
Pod Vodárenskou věží 4, 182 08 Praha 8, Czech Republic

Abstract

Compositional model theory serves as an alternative approach to multidimensional probability distribution representation and processing. Every compositional model over a finite non-empty set of variables N is uniquely defined by its generating sequence – an ordered set of low-dimensional probability distributions. A generating sequence structure induces a system of conditional independence statements over N valid for every multidimensional distribution represented by a compositional model with this structure.

The *equivalence problem* is how to characterise whether all independence statements induced by structure $\mathcal{P}$ are induced by a second structure $\mathcal{P}'$ and vice versa. This problem can be solved in several ways. A partial solution of the so-called *direct characterisation* of an equivalence problem is represented here. We deduce and describe three properties of equivalent structures invariant in a class of equivalent structures. We call them *characteristic properties* of equivalence.

Keywords: Equivalence problem, compositional model, persegram, characteristic properties

1. Introduction

The ability to represent and process multidimensional probability distributions is a necessary condition for the application of probabilistic methods in artificial intelligence. Among the most popular approaches are the methods based on graphical Markov models, e.g., Bayesian networks. The compositional models (see [1] or [4] for example) represent an alternative approach to graphical Markov models.

A Bayesian network may be defined as a multidimensional distribution factorising with respect to an acyclic directed graph, or it may alternatively be defined by its graph and an appropriate system of low-dimensional *conditional distributions*. Similarly, a compositional model is defined as a multidimensional distribution assembled from a sequence of low-dimensional *unconditional distributions*, with the aid of an *operator of composition*. We call the sequence of low-dimensional distributions a *generating sequence* of the compositional model. The main advantage of both approaches lies in the fact that low-dimensional distributions could easily be stored in a computer memory. However, computations on a multidimensional distribution that is split into many pieces may be exceptionally complicated.

There are two main advantages to using compositional models as compared to Bayesian networks. First, compositional models explicitly express some marginals, whose computation in a Bayesian network may be demanding. Secondly, no auxiliary graphical tool, such as a directed acyclic graph, is required in compositional models.

As stated above, a compositional model is *composed* from an ordered system of low-dimensional distributions – the so-called generating sequence. The binary *operator of composition* used during this process is basically a normalised product of its parameters designed to create a probability distribution over the union of variables for which the input distributions are defined. (See Definition 2.1.) While the model is put together, *(un)conditional independencies* are simultaneously introduced by *the structure of the generating sequence*. For example, for a two-dimensional distribution composed from two one-dimensional ones, the respective variables are independent.

Email address: `velorex@utia.cas.cz` (Václav Kratochvíl)

For the sake of clarity, a *structure of a compositional model* (a structure of its generating sequence in particular) can be visualised by a tool called a *persegram*, and one can read the induced independencies directly using this tool. We say that every structure (or its corresponding persegram) induces an *independence model* – a list of (un)conditional independence statements.

The *equivalence problem* is how to characterise whether an independence model induced by one structure is identical with an independence model induced by another structure, and vice versa. Structures inducing the same independence model are said to be *equivalent*. One can find two different approaches to solve this problem in other probability models. First is the so called *direct characterisation*, which is based on several characteristic structure properties sufficient to guarantee the equivalence. Second, a group of local transformations preserving the independence model can be found and two structures are equivalent if there is a sequence of these transformations from one to the next.

This paper puts forth two major contributions in this area. First, we derive two characteristic properties of equivalent structures which can later be used for direct characterisation of an equivalence problem. The second contribution is presentation of a very special subset of the relevant structure – the so called *structure core*. It results from a very new approach to sets of variables, where we distinguish whether the set is *trivial or non-trivial* in this structure. The *structure core* corresponds to the so-called *reduced persegram* published in [6]. The local transformations preserving an induced independence model (published in [7]) now seem to be a very logical consequence of these properties.

2. Notation

Throughout the paper the symbol N will denote a non-empty finite set of finite-valued *variables*. The symbols K, U, V, W, Z will be used for subsets of N. $|U|$ will denote the number of elements in U, that is, its *cardinality*. Symbols u, v, w, x, y, z denote variables as well as singletons $\{u\}, \{v\}, \{w\}, \{x\}, \{y\}, \{z\}$. Two set inclusion symbols are used thorough the paper, namely $\subset$ and $\subseteq$. Whereas the symbol $\subseteq$ represents the usual (non-strict) case of inclusion, the symbol $\subset$ is used for strict inclusion only. That means if $U \subset V$ then $V \setminus U \neq \emptyset$.

All probability distributions of the variables from N will be denoted by Greek letters (usually π); thus for $K \subseteq N$, we consider a distribution $\pi(K)$ which is defined on variables K. If we work with several distributions, we distinguish between them by indices. For a probability distribution $\pi(K)$ and $U \subseteq K$ we denote the respective *marginal distribution* $\pi(U)$ or $\pi^{\downarrow U}$.

For a probability distribution $\pi(N)$ and three disjoint subsets $U, V, Z \subseteq N$ such that $U \neq \emptyset \neq V$, we say that sets of variables U and V are *conditionally independent* given Z in π (in symbol $U \perp\!\!\!\perp V | Z[\pi]$) if

$$\pi^{\downarrow U \cup V \cup Z}(x) \cdot \pi^{\downarrow Z}(x) = \pi^{\downarrow U \cup Z}(x) \cdot \pi^{\downarrow V \cup Z}(x).$$

for all $x \in \times_{j \in U \cup V \cup Z} \mathbf{X}_j$. Observe that, if $Z = \emptyset$, then the conditional independence coincides with unconditional independence. The unconditional independence of variable sets U and V in π is denoted by $U \perp\!\!\!\perp V[\pi]$

The keystone of Compositional Models is an operator of composition $\triangleright$. It is used to compose low-dimensional distributions to get a distribution of a higher dimension. The composition is described in the following definition.

Definition 2.1. *For two arbitrary distributions $\pi_1(U)$ and $\pi_2(V)$ their* composition *is given by the formula*

$$\pi_1(U) \triangleright \pi_2(V) = \frac{\pi_1(U)\pi_2(V)}{\pi_2(U \cap V)}$$

if $\pi_1(U \cap V) \ll \pi_2(U \cap V)$, otherwise the composition remains undefined.

The symbol $\pi_1(K) \ll \pi_2(K)$ means that $\pi_1(K)$ is dominated *by $\pi_2(K)$, which in its turn means (in the considered finite setting) $\forall x \in \times_{j \in K} \mathbf{X}_j; (\pi_2(x) = 0 \implies \pi_1(x) = 0)$. Moreover, if for any $x \in \times_{j \in U \cap V} \mathbf{X}_j$ $\pi_2(x) = 0$, then by dominance $\pi_1(U \cap V) \ll \pi_2(U \cap V)$ there is a product of two zeros in the numerator and we take $\frac{0 \cdot 0}{0} = 0$*

The result of the composition (if defined) is a new distribution. We can iteratively repeat the process of composition to obtain a multidimensional distribution. That is why the multidimensional distribution is called a *compositional model*. Regarding the fact that the operator $\triangleright$ is neither commutative nor associative, we always apply the operator from left to right; e.g.,

$$\pi_1(K_1) \triangleright \pi_2(K_2) \triangleright \ldots \triangleright \pi_n(K_n) = (\ldots(\pi_1(K_1) \triangleright \pi_2(K_2)) \triangleright \ldots) \triangleright \pi_n(K_n).$$

Therefore, in order to construct such a model it is sufficient to determine a sequence of low-dimensional distributions $\pi_1, \pi_2, \ldots, \pi_n$ – we call it a *generating sequence* .

From now on, we consider a generating sequence $\pi_1(K_1)$, $\pi_2(K_2), \ldots,$ $\pi_n(K_n)$ such that $\pi_1(K_1) \triangleright$ $\pi_2(K_2) \triangleright \ldots \triangleright \pi_n(K_n)$ is defined. Therefore, whenever distribution π_i is used, we assume it is defined for variables K_i. A sequence of sets K_1, $K_2, \ldots,$ K_n is called *model structure* and it is denoted by $\mathcal{P}$. If not specified otherwise, $\mathcal{P} = K_1, \ldots,$ K_n where $(K_1 \cup \ldots \cup K_n) = N$, and we say that $\mathcal{P}$ is defined over N and $K_i \in \mathcal{P}$ for every $i \in \{1, \ldots, n\}$. One may denote K_i as $K_i^{\mathcal{P}}$ to emphasise that $K_i \in \mathcal{P}$. In addition, each set K_i can be divided into two disjoint parts with respect to model structure. We denote them $R(K_i)$ and $S(K_i)$:

$$\begin{aligned} R(K_i) &= K_i \backslash (K_1 \cup \ldots \cup K_{i-1}) \\ S(K_i) &= K_i \cap (K_1 \cup \ldots \cup K_{i-1}) \end{aligned}$$

It has the following meaning: $R(K_i)$ denotes the variables first occurring in the sequence (meaning from left to right). $S(K_i)$ denotes the variables which have already been used. Observe that $K_i = R(K_i) \cup S(K_i)$. $|\mathcal{P}|$ denotes the number of sets in the structure, i.e., $|\mathcal{P}| = n$ for $\mathcal{P} = K_1, \ldots, K_n$.

As stated in the introduction, while a model is put together, a system of (un)conditional independencies is simultaneously introduced by the structure of the generating sequence.

Example 2.2. *Let $\{u, v\} = N$, $u \neq v$. $\pi_1(u), \pi_2(v)$ is a generating sequence of a compositional model $\pi_1 \triangleright \pi_2$. Then $u \perp\!\!\!\perp v[\pi_1 \triangleright \pi_2]$. Indeed, by applying the operator of composition one gets*

$$\pi_1(u) \triangleright \pi_2(v) = \frac{\pi_1(u)\pi_2(v)}{\pi_2(\emptyset)} = \pi_1(u)\pi_2(v),$$

which corresponds to the definition of independence of variables u and v.

Similarly, assume $\{u, v, w\} = N$ are three distinct variables $\pi_1(u, w)$, and $\pi_2(v, w)$ is a generating sequence of a compositional model $\pi_1 \triangleright \pi_2$. Using Definition 2.1 we get

$$\pi_1(u, w) \triangleright \pi_2(v, w) = \frac{\pi_1(u, w)\pi_2(v, w)}{\pi_2(w)}.$$

Then $u \perp\!\!\!\perp v|w[\pi_1 \triangleright \pi_2]$ by the definition of conditional independence.

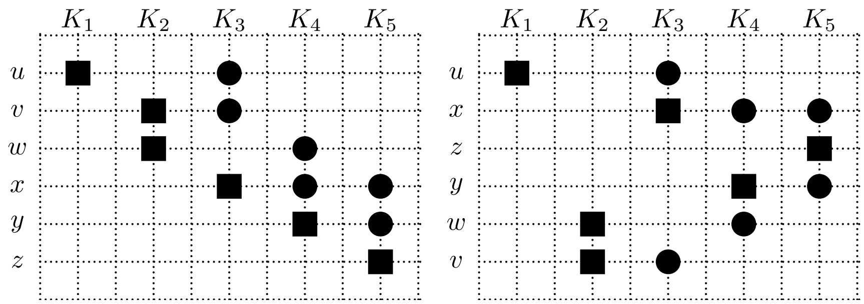
The more complex the model structure is, the more difficult the seeking of induced independencies is. Let us note that the set of independencies induced by a structure is valid for any compositional model with this structure regardless the generating distributions' properties. Obviously, one can read induced independencies directly from the model structure. To increase the lucidity and readability of this text, we have decided to use a specific visualisation of the structure, and we present the procedure for reading induced independencies using this tool.

2.1. Persegrams

It is well-known that one can read conditional independence relations of a Bayesian network from its graph. A similar technique has been developed for compositional models. An appropriate tool for this is a *persegram* – a visualisation tool of the model structure.

Definition 2.3. Persegram *of a structure $\mathcal{P} = K_1, K_2, \ldots, K_n$ is a table in which rows correspond to variables from $K_1 \cup K_2 \cup \ldots \cup K_n$ (in an arbitrary order) and columns to sets of variables K_i for all $i \in \{1, \ldots, n\}$; ordering of the columns corresponds to the structure ordering. A position in the table is marked if the respective set contains the corresponding variable. Markers for the first occurrence of each variable (i.e., the leftmost markers in rows) are* box-markers *and for other occurrences there are* bullets*.*

Example 2.4. *Let $\mathcal{P} = K_1, \ldots, K_5$ be structure of a compositional model such that $K_1 = \{u\}, K_2 = \{v, w\}, K_3 = \{u, v, x\}, K_4 = \{w, x, y\}, K_5 = \{x, y, z\}$. Since the row ordering is not specified in Definition 2.3, the corresponding persegram can be visualised not only as in Figure 1a, but also in many other ways. See another persegram in Figure 1b.*



(a) (b)

Figure 1: Different persegrams belonging to one model structure $\mathcal{P}$

Observe that bullets in i-th column correspond to $S(K_i^{\mathcal{P}})$ and box-markers to $R(K_i^{\mathcal{P}})$.

To be able to simply handle characteristic properties of the respective structures and persegrams, we introduce a function

$$]\cdot[_{\mathcal{P}}\colon N \to \{1,\ldots,n\}$$

such that $]u[_{\mathcal{P}}$ returns the index of set $K_i^{\mathcal{P}}$ with the first appearance of u in $\mathcal{P}$ for every variable $u \in (K_1^{\mathcal{P}} \cup \ldots \cup K_n^{\mathcal{P}})$. Hence, $]u[_{\mathcal{P}}= min\{i : u \in K_i^{\mathcal{P}}\}$. Due to the previously established notation, it can be said that $K^{\mathcal{P}}_{]u[_{\mathcal{P}}}$ is a column $K_i^{\mathcal{P}}$ where $u \in R(K_i^{\mathcal{P}})$, i.e., $]u[_{\mathcal{P}}= i : u \in R(K_i^{\mathcal{P}})$. The symbol $\mathcal{P}$ may be omitted in $]u[_{\mathcal{P}}$ if the context is clear.

Example 2.5. *Take the model structure $\mathcal{P} = K_1,\ldots,K_5$ from Example 2.4. One can read the following properties from both of its persegrams in Figure 1:* $]u[_{\mathcal{P}}= 1$*,* $]v[_{\mathcal{P}}= 2$*,* $]w[_{\mathcal{P}}= 2$*,* $]x[_{\mathcal{P}}= 3$*,* $]y[_{\mathcal{P}}= 4$*,* $]z[_{\mathcal{P}}= 5$*, and*

$R(K_1^{\mathcal{P}}) = \{u\}$ $\quad S(K_1^{\mathcal{P}}) = \emptyset$
$R(K_2^{\mathcal{P}}) = \{v, w\}$ $\quad S(K_2^{\mathcal{P}}) = \emptyset$
$R(K_3^{\mathcal{P}}) = \{x\}$ $\quad S(K_3^{\mathcal{P}}) = \{u, v\}$
$R(K_4^{\mathcal{P}}) = \{y\}$ $\quad S(K_4^{\mathcal{P}}) = \{w, x\}$
$R(K_5^{\mathcal{P}}) = \{z\}$ $\quad S(K_5^{\mathcal{P}}) = \{x, y\}$

Definition 2.6. *For arbitrary variables $u, v \in N$ and structure $\mathcal{P}$ over N we introduce a binary relation $u \preceq_{\mathcal{P}} v$ such that $u \preceq_{\mathcal{P}} v$ if and only if $]u[_{\mathcal{P}}\leq]v[_{\mathcal{P}}$. Moreover, we introduce the relation $\prec_{\mathcal{P}}$: $u \prec_{\mathcal{P}} v \Leftrightarrow]u[_{\mathcal{P}}<]v[_{\mathcal{P}}$.*

The following convention will be used throughout the paper: Given a structure $\mathcal{P}$ over N, set $U \subseteq N$ and variable $v \in N$, the term $U \prec_{\mathcal{P}} v$ denotes that $u \prec_{\mathcal{P}} v$ for all $u \in U$. The symbol $\mathcal{P}$ may be omitted if the context is clear.

Example 2.7. *Let $K_1,\ldots,K_5$ be the same model structure as in Example 2.4 again. According to the former definition one can see that $u \prec v \preceq w \prec x \prec y \prec z$ in both persegrams from Figure 1.*

2.2. Induced models

In this section we shall demonstrate how to read induced conditional independence relations from a persegram representing a structure of a compositional model. Such independencies are indicated by the absence of a *trail connecting relevant markers and avoiding others* which is defined below.

Definition 2.8. *A sequence of markers $m_0,\ldots,m_t$ of a persegram corresponding to structure $\mathcal{P}$ is called a Z-*avoiding trail *($Z \subseteq K_1^{\mathcal{P}} \cup \ldots \cup K_{|\mathcal{P}|}^{\mathcal{P}}$) that connects m_0 and m_t if it meets the following five conditions:*

0. *m_0 and m_t do not correspond to a variable from Z*
1. *for each $s = 1,\ldots,t$ a couple (m_{s-1}, m_s) is either in the same row (i.e., horizontal connection) or in the same column (vertical connection);*
2. *each vertical connection must be adjacent to a box-marker;*
3. *no horizontal connection corresponds to a variable from Z;*

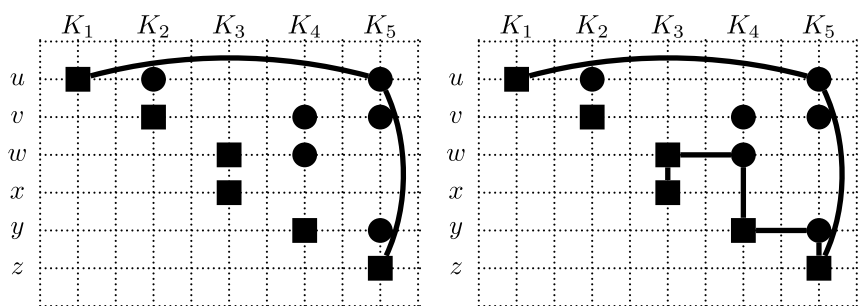
4. *vertical and horizontal connections regularly alternate with the following possible exception: at most two vertical connections may be in direct succession if their common adjacent marker is a box-marker of a variable from Z;*

If a Z-avoiding trail connects two box-markers corresponding to variables u and v, we say that these variables are connected by a Z-avoiding trail. *This situation will be denoted by* $u \not\perp\!\!\!\perp v|Z[\mathcal{P}]$.

By investigating Definition 2.8 further, the reader will find that no condition of the definition is dependent on the order of rows in the considered persegram. That would not be appropriate either, because all persegrams representing the structure of a generating sequence are equivalent regardless of the row ordering. (See the definition of persegram - Definition 2.3). Then the system of Z-avoiding trails induced by a persegram can be obtained by any other persegram of the considered structure. In the sense of the previous definition, all persegrams corresponding to $\mathcal{P}$ are equivalent.

Example 2.9. *Consider a persegram visualising a structure* $\mathcal{P}$ *as it is depicted in Figure 2. There is a sequence of markers in each part of it. In order to illustrate vertical and horizontal connections and to highlight the ordering, each two consecutive markers are connected with a line.*

There is a sequence of markers $[K_1, u], [K_5, u], [K_5, z]$ *in Figure 2a. Considering* $Z = \emptyset$*, it forms a Z-avoiding trail connecting u and z. However, considering Definition 2.8, this sequence avoids many other variables and Z may have various content. In fact, Z can be any subset of* $\{v, w, x, y\}$*. Hence,* $u \not\perp\!\!\!\perp v|Z[\mathcal{P}]$ *for any* $Z \subseteq \{v, w, x, y\}$.



(a) $\mathcal{P} : u \not\perp\!\!\!\perp z|\emptyset,\ u \not\perp\!\!\!\perp z|\{v, w, x, y\}$ (b) $\mathcal{P} : u \not\perp\!\!\!\perp x|\{z\},\ u \not\perp\!\!\!\perp x|\{v, z\}$

Figure 2: Different trails connecting u with some other variables

Similarly, the sequence of markers $[K_1, u]$, $[K_5, u]$, $[K_5, z]$, $[K_5, y]$, $[K_4, y]$, $[K_4, w]$, $[K_3, w]$, $[K_3, x]$ *from Figure 2b is a* $\{z\}$*-avoiding trail. Contrary to Figure 2a, one cannot replace z by any other variable. Otherwise, the 4th condition from Definition 2.8 would be corrupted. However, the trail depicted in Figure 2b is* $\{v, z\}$*-avoiding too.*

With the help of Z-avoiding trails, the so-called *(un)conditional (in)dependencies induced by a persegram* are introduced.

Definition 2.10. *Consider a persegram corresponding to a structure* $\mathcal{P}$ *over N and three disjoint subsets* $U, V, Z \subset N$ *such that* $U \neq \emptyset \neq V$*. The sets of variables U and V are* conditionally independent given Z in $\mathcal{P}$ *(in symbol* $U \perp\!\!\!\perp V|Z[\mathcal{P}]$*), if no* $u \in U$ *is connected with a* $v \in V$ *by a Z-avoiding trail. Otherwise U and V are* conditionally dependent given by Z in $\mathcal{P}$*, written* $U \not\perp\!\!\!\perp V|Z[\mathcal{P}]$.

The induced independence model $\mathcal{I}(\mathcal{P})$ *and the* induced dependence model $\mathcal{D}(\mathcal{P})$ *of structure* $\mathcal{P}$ *are defined as follows:*

$$\mathcal{I}(\mathcal{P}) = \{\langle U, V|Z\rangle \in \mathcal{T}(N); U \perp\!\!\!\perp V|Z[\mathcal{P}]\}$$

$$\mathcal{D}(\mathcal{P}) = \{\langle U, V|Z\rangle \in \mathcal{T}(N); U \not\perp\!\!\!\perp V|Z[\mathcal{P}]\},$$

where the symbol $\mathcal{T}(N)$ *denotes the class of all disjoint triplets over N:*

$$\mathcal{T}(N) = \{\langle U, V|Z\rangle : U, V, Z \subseteq N, U \neq \emptyset \neq V, U \cap V = V \cap Z = Z \cap U = \emptyset\}$$

The concept of induced *(in)dependencies* lives up to expectations that there is a parallel between this and independencies valid in any compositional model with the same structure. The connection between independence read from a compositional model and from its persegram is elucidated by the following theorem. The proof can be found in [2].

Theorem 2.11. *Consider a generating sequence $\pi_1(K_1), \ldots, \pi_n(K_n)$, the corresponding structure $\mathcal{P}$, and three disjoint subsets $U, V, Z \subseteq K_1 \cup \ldots \cup K_n$ such that $U \neq \emptyset \neq V$. Then:*

$$U \perp\!\!\!\perp V | Z[\mathcal{P}] \Rightarrow U \perp\!\!\!\perp V | Z[\pi_1 \triangleright \ldots \triangleright \pi_n].$$

It is important to realise that (analogously to the situation when Bayesian networks or decomposable models are considered) one can be sure about the validity of the indicated independence relations for any distribution which is represented by a compositional model with the given persegram (structure).

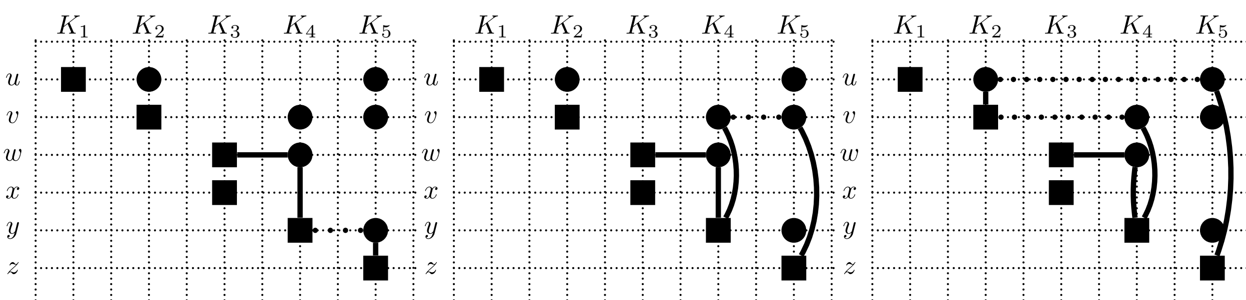
2.3. Other preliminaries

A trivial fact follows from Definition 2.8. It concerns variables appearing for the first time in the last column . Before we introduce this fact in the form of a lemma, let us illustrate it with the help of the following example.

Example 2.12. *Consider the persegram from Figure 3. I would like to show that there is no Z-avoiding trail connecting $z \in R(K_5)$ (first appearing in the last column) with $w \notin K_5$ (not belonging to the last column) for $Z = \{u, v, y\}$. Let us try to construct such a sequence of markers forming a Z-avoiding trail.*

Three different sequences of markers are depicted in Figure 3. Let us summarise requirements necessary for these sequences to be Z-avoiding trails:

- *Consider the sequence of markers highlighted in Figure 3a: By the 3rd condition of Definition 2.8 (no horizontal connection corresponds to a variable from Z), Z must not contain a variable y ($y \notin Z$).*
- *Figure 3b: Similarly, $v \notin Z$ for the same reasons.*
- *Figure 3c: $u, v \notin Z$.*



(a) $\mathcal{P} : y \in Z$ (b) $\mathcal{P} : v \in Z$ (c) $\mathcal{P} : u, v \in Z$

Figure 3: Different trails violating 3rd condition of Definition 2.8 if $Z = \{u, v, y\}$.

Combining the restrictions on Z together, one gets the following corollary: By choosing $Z = S(K_5) = \{u, v, y\}$, none of the above-discussed sequences forms a Z-avoiding trail since each of them contains a horizontal connection corresponding to a variable from $S(K^{\mathcal{P}}_{|\mathcal{P}|})$. These horizontal connections violating the 3rd condition of Definition 2.8 are drawn by dotted lines. Since there is no other possible $S(K_5)$-avoiding trail between w and z, $w \perp\!\!\!\perp z | S(K_5)$ holds due to Definition 2.10.

Lemma 2.13. *Consider a structure $\mathcal{P} = K_1, \ldots, K_n$ and distinct variables $u, v \in (K_1 \cup \ldots \cup K_n)$ such that $u \in R(K_n)$ and $v \notin K_n$. Then $u \perp\!\!\!\perp v | S(K_n) [\mathcal{P}]$.*

Proof. Consider a persegram of $\mathcal{P}$. Since u belongs to the last column of $\mathcal{P}$ ($u \in K^{\mathcal{P}}_{|\mathcal{P}|}$), every trail from u has to begin with a vertical connection in K_n to a marker corresponding to a variable from $S(K_n)$ (otherwise, in a case where the vertical connection connects two variables from $R(K_n)$, the horizontal and vertical connections could not regularly alternate). However, no $S(K_n)$-avoiding trail may contain a horizontal connection corresponding to a variable from $S(K_n)$, and such a trail must not contain any marker out of the last column. Since $u \notin K_n$, a trail representing $u \not\perp v | S(K_n)$ cannot exist; therefore, $u \perp v | S(K_n) [\mathcal{P}]$ by Definition 2.10. □

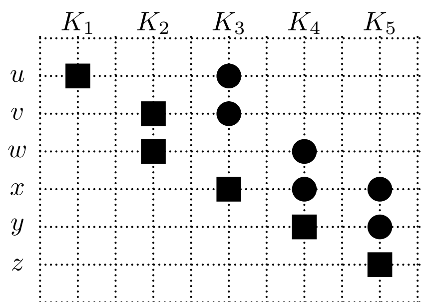
To simplify the following, we introduce the concept of the *substructure induced by a set of variables*. Unlike the subgraph which contains exactly those variables that induce it, the substructure is usually defined for some superset.

Definition 2.14. *A* substructure *of a structure $\mathcal{P} = K_1, \ldots, K_n$ induced by a set $U \subseteq (K_1 \cup \ldots \cup K_n)$ is its minimal left part containing all variables U. $\mathcal{P}[U] = K_1, \ldots, K_{max\{\}u[:u \in U\}}$*

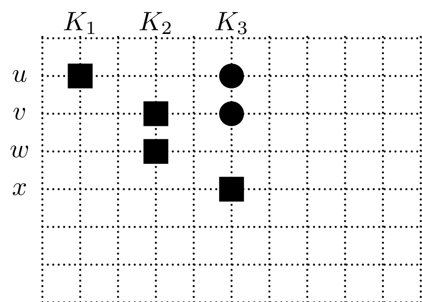
Persegram of $\mathcal{P}[U]$ is created from persegram of $\mathcal{P}$ by removing columns to the right of the one with the farthest right box-marker corresponding to a variable from U.

Remark 2.15. *Observe that, given U and $Z \subset U$, any sequence of markers forming a Z-avoiding trail in a persegram of $\mathcal{P}[U]$ forms a Z-avoiding trail in a persegram of $\mathcal{P}$.*

Example 2.16. *Consider a structure $\mathcal{P} = K_1, \ldots, K_5$ from Example 2.4 again. Its corresponding persegram is in Figure 4a. Suppose $U = \{u, x\}$ holds. One can then find a persegram of the induced persegram $\mathcal{P}[U]$ in Figure 4b. Observe that $\mathcal{P}[U]$ is defined not only over $\{u, x\}$, but also over $\{v, w\}$.*



(a) Persegram of $\mathcal{P}$



(b) Persegram of substructure

Figure 4: Visualisation of a structure and its substructure which is induced by $\{u, x\}$

The concept of an induced substructure brings one very important advantage. Searching of Z-avoiding trails connecting u with v in a persegram may be restricted to a persegram of its substructure induced only by $\{u, v\} \cup Z$.

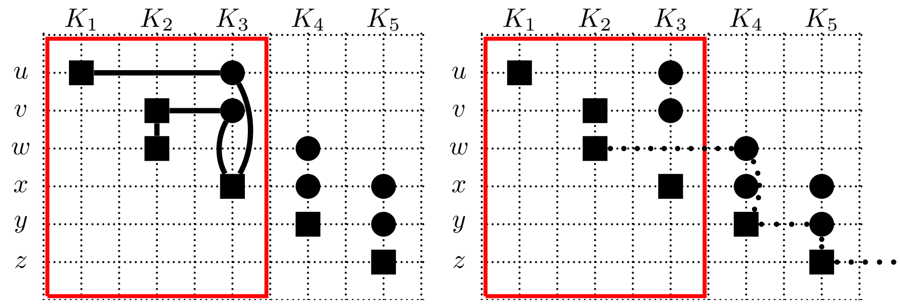
Lemma 2.17. *Consider a persegram of structure $\mathcal{P}$ over N, $u, v \in N$, and $Z \subseteq N \setminus \{u, v\}$. If $u \not\perp v | Z[\mathcal{P}]$, then all Z-avoiding trails connecting u with v are in the persegram of its substructure $\mathcal{P}[\{u, v\} \cup Z]$.*

Proof. Assume that there is a Z-avoiding trail representing $u \not\perp v | Z[\mathcal{P}]$ containing markers out of the area defined by $\mathcal{P}[\{u, v\} \cup Z]$. We show that if Z-avoiding trail from u leaves the area defined by $\mathcal{P}[\{u, v\} \cup Z]$, then it cannot end up in v which contradicts the assumption. To understand our way of thinking, the reader should have a careful look at Figure 5b during the procedure.

Assume that τ is a Z-avoiding trail representing $u \not\perp v | Z[P]$ with a marker out of $\mathcal{P}[\{u, v\} \cup Z]$. I.e. $\tau = m_0, \ldots, m_t$ is a sequence of markers where m_0 corresponds to u. Let m_i be the first marker in the sequence τ such that it comes out of the part of persegram corresponding to $\mathcal{P}[\{u, v\} \cup Z]$. Since it is the first marker in such a column, a horizontal connection had to be used between m_{i-1} and m_i and therefore m_i has to be a bullet. By Definition 2.8, the trail now has to continue with a vertical connection to a box-marker. Since this box-marker cannot correspond to any variable from Z (it is out of $\mathcal{P}[\{u, v\} \cup Z]$), one has to continue with a horizontal connection (by definition, to the right of the box-marker (first marker in the row) – there is nothing on the left in the same row) to a bullet. Then we again make a vertical connection to a box-marker which does not correspond to any variable from Z, etc. From such a trail τ, there is no return to v. Therefore such a trail cannot exist, which contradicts the assumption. □

Example 2.18. *Let us illustrate the idea of proving Lemma 2.17. Consider the persegram from Figure 5, its corresponding structure $\mathcal{P}$, and $Z = \{x\}$. I am going to show that there is only one $\{x\}$-avoiding trail representing $u \not\perp\!\!\!\perp w|x$. One can find it in the area corresponding to $\mathcal{P}[\{u, w, x\}]$ in Figure 5a.*

Let us try to create an x-avoiding trail from w to u containing markers out of the highlighted part corresponding to $\mathcal{P}[\{u, w, x\}]$. Such an experiment is depicted by the dotted line in Figure 5b.



(a) The only x-avoiding trail connecting u with w. It is located in the area induced by $\{u, w, x\}$

(b) An attempt to create a z-avoiding trail outside the area corresponding to induced substructure.

Figure 5: Illustration of Lemma 2.17

Let us start in the box-marker (K_2, w) and continue out of $\mathcal{P}[\{u, w, x\}]$ into (K_4, w). To satisfy Definition 2.8 of a Z-avoiding trail, one has to continue with a vertical connection to a box-marker. (The only possible box-marker is (K_4, y)). Since $y \notin Z$, then by the 4th condition of Definition 2.8 one has to continue with a horizontal connection (to the right – there is nothing left of any box-marker), etc. Since there is no box-marker corresponding to Z outside of $\mathcal{P}[\{u, w, x\}]$, the trail moves away from u. Since there is no return for such a trail, it cannot exist.

Lemma 2.17 basically means that, if we are interested in relation $u \perp\!\!\!\perp v|Z[\mathcal{P}]$, we may focus only on the subpersegram $\mathcal{P}[\{u, v\} \cup Z]$. This observation is summarised in the following corollary.

Corollary 2.19. *Let $\mathcal{P}$ be a persegram over N and $u, v \in N$, $Z \subseteq N \setminus \{u, v\}$. Then*

$$u \perp\!\!\!\perp v|Z\ [\mathcal{P}[\{u, v\} \cup Z]] \Leftrightarrow u \perp\!\!\!\perp v|Z\ [\mathcal{P}].$$

Proof. The proof is a trivial consequence of Lemma 2.17 and Remark 2.15. □

3. Equivalence problem

By the equivalence problem we understand how to recognise whether two given structures $\mathcal{P}, \mathcal{P}'$ over the same set of variables N induce the same independence model $(\mathcal{I}(\mathcal{P}) = \mathcal{I}(\mathcal{P}'))$. A very readable overview of the solution to this problem using Bayesian networks may be found in [3].

It is of special importance to have a simple rule to recognise that two structures are equivalent in this sense (the notion of a rule simplicity may differ when considering whether people or a computer will use it), and an easy way to convert $\mathcal{P}$ into $\mathcal{P}'$ in terms of some elementary operations on structures. These issues are addressed in [5], [6] and [7]. Another very important aspect is the ability to generate all structures which are equivalent to a given structure.

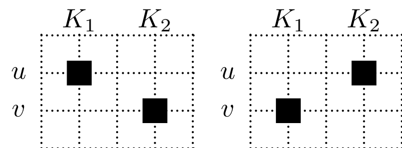
We only focus on one part of the equivalence problem in this paper. We introduce and describe two properties of a model structure which are characteristics of a class of equivalent structures. This means that they are necessary to guarantee the equivalence of different structures. They include the so-called *connection set* and *F-condition set*. However, as discussed at the end of this section, the *connection set* is not as easily verifiable in cases involving more complex structures; therefore, we introduce another property based on the connection set – the so-called *core inclusion*, which has very interesting consequences.

Definition 3.1. *Structures $\mathcal{P}, \mathcal{P}'$ (over the same variable set N) are called* independence equivalent*, if they induce the same independence model $\mathcal{I}(\mathcal{P}) = \mathcal{I}(\mathcal{P}')$.*

Remark 3.2. *One may easily see that the above-mentioned definition could be formulated using a dependence model instead. Structures $\mathcal{P}, \mathcal{P}'$ (over the same variable set N) are independence equivalent iff $\mathcal{D}(\mathcal{P}) = \mathcal{D}(\mathcal{P}')$. This alternative is primarily used in most proofs.*

Example 3.3.

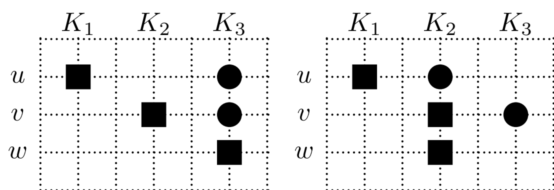
1. *Consider two simple structures $\mathcal{P}_1, \mathcal{P}'_1$ over $\{u, v\}$ as they are depicted in Figure 6 by corresponding persegrams. Since there is no possible vertical connection in both persegrams, there can be no Z-avoiding trail for any Z in these persegrams. Therefore $u \perp\!\!\!\perp v|\emptyset$ in both $\mathcal{P}_1$ and $\mathcal{P}'_1$. Hence $\mathcal{I}(\mathcal{P}_1) = \mathcal{I}(\mathcal{P}'_1) = \{\langle u, v|\emptyset\rangle\}$. The corresponding structures are independence equivalent.*



(a) $\mathcal{P}_1$ (b) $\mathcal{P}'_1$

Figure 6: Persegrams of two equivalent structures

2. *On the other hand, structures with the same sets ($K_i \in \mathcal{P}_2 \Leftrightarrow K_i \in \mathcal{P}'_2$) in a different ordering are not equivalent. Let $N = \{u, v, w\}$ and consider the following structures $\mathcal{P}_2, \mathcal{P}'_2$ visualised in Figure 7. Observe that $u \perp\!\!\!\perp v|\emptyset[\mathcal{P}_2]$ but $u \not\perp\!\!\!\perp v|\emptyset[\mathcal{P}'_2]$. On the contrary, $u \not\perp\!\!\!\perp v|w[\mathcal{P}_2]$ and $u \not\perp\!\!\!\perp v|w[\mathcal{P}'_2]$. The set ordering is important.*



(a) $\mathcal{P}_2$ (b) $\mathcal{P}'_2$

Figure 7: Persegrams of two non-equivalent structures

Recall that each Z-avoiding trail contains one or several vertical connections. However, contrary to the persegram from Figure 7b, there is no possible vertical connection between markers corresponding to variables u, v in the persegram from Figure 7a. That is why these structures are not equivalent. One of the characteristic properties is based on this observation.

3.1. Characteristic properties

Now, step by step, we deduce two structural properties necessary for independence equivalence of the respective structures: the *connection set* and the so-called *F-condition set* necessary for independence equivalence of the underlying structures. The proof of sufficiency of these properties is not included in this paper.

3.1.1. Connection set

Two structures are equivalent if and only if they induce the same dependence models. The dependence relation is represented by a Z-avoiding trail in the corresponding persegram. Thus, in case of two equivalent structures, one should be able to create the same set of Z-avoiding trails including the elementary ones that are composed only of two markers – one vertical connection.

It turns out that the set of vertical connections is just the characteristic property of a class of equivalent persegrams.

Definition 3.4. *Consider a structure $\mathcal{P} = K_1, \ldots K_n$ and two distinct variables $u, v \in (K_1 \cup \ldots \cup K_n)$. We say that u, v are* connected *in $\mathcal{P}$ ($u \leftrightarrow_{\mathcal{P}} v$) iff $u \in K^{\mathcal{P}}_{]v[}$ or $v \in K^{\mathcal{P}}_{]u[}$. The set of all pairs $\mathcal{E}(\mathcal{P}) = \{\langle u, v\rangle : u, v \in N, u \leftrightarrow_{\mathcal{P}} v\}$ is called a* connection set *of $\mathcal{P}$.*

Remark 3.5. *The previous definition basically means that u, v are* connected *in $\mathcal{P}$ iff there is a column in its persegram containing markers of both variables and at least one of them is a box-marker. It means that $u \leftrightarrow v$ corresponds to vertical connection from Definition 2.8.*

The following convention will be used throughout the paper: Given variable $w \in N$, $U \subseteq N \setminus \{w\}$ and a structure $\mathcal{P}$ over N, the term $U \leftrightarrow_{\mathcal{P}} w$ denotes that $u \leftrightarrow_{\mathcal{P}} w$ for every $u \in U$. The symbol $\mathcal{P}$ may be omitted if the context is clear.

For purposes of the following text, one should realise that when $u \leftrightarrow_{\mathcal{P}} v$, there is an obvious parallel between ordering of variables u, v and content of respective columns $K_{]u[}, K_{]v[}$. It is summarised in the following trivial lemma.

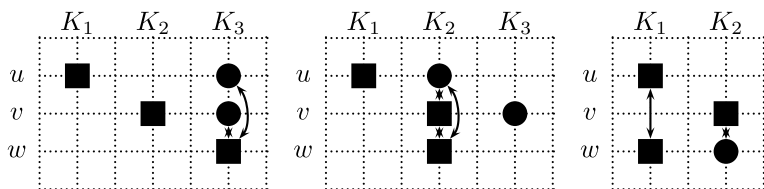
Lemma 3.6. *Let $\mathcal{P} = K_1, \ldots, K_n$ be a structure and $u, v \in (K_1 \cup \ldots \cup K_n)$ two distinct variables. Then*

$$u \preceq_{\mathcal{P}} v \text{ and } u \leftrightarrow_{\mathcal{P}} v \Leftrightarrow u \in K_{]v[}^{\mathcal{P}}.$$

Proof. The lemma is a trivial consequence of Definition 3.4. □

Observe that $u \in S(K_{]v[}^{\mathcal{P}})$ in the previous lemma in case of strict version $u \prec_{\mathcal{P}} v$.

Example 3.7. *Consider three different structures $\mathcal{P}_2, \mathcal{P}'_2, \mathcal{P}_3$ depicted in Figure 8.*



(a) $\mathcal{P}_2$ (b) $\mathcal{P}'_2$ (c) $\mathcal{P}_3$

Figure 8: Connections in different persegrams

One can read the following relations using persegrams from Figure 8:

$\mathcal{P}_2$:	$\{u, v\} \leftrightarrow_{\mathcal{P}_2} w$	$\mathcal{E}(\mathcal{P}_2) = \{\langle u, w\rangle, \langle v, w\rangle\}$
$\mathcal{P}'_2$:	$u \leftrightarrow_{\mathcal{P}'_2} w, v \leftrightarrow_{\mathcal{P}'_2} w, u \leftrightarrow_{\mathcal{P}'_2} v$	$\mathcal{E}(\mathcal{P}'_2) = \{\langle u, w\rangle, \langle v, w\rangle, \langle u, v\rangle\}$
$\mathcal{P}_3$:	$u \leftrightarrow_{\mathcal{P}_3} w, v \leftrightarrow_{\mathcal{P}_3} w$	$\mathcal{E}(\mathcal{P}_3) = \{\langle u, w\rangle, \langle v, w\rangle\} = \mathcal{E}(\mathcal{P}_2)$

As previously stated, the connection $u \leftrightarrow v$ corresponds to the existence of a vertical connection between markers corresponding to u, v. Therefore, if there is a connection between two variables, then there is a simple trail connecting the corresponding variables. Since the trail contains no other markers, it is Z-avoiding for any Z such that $Z \subseteq N \setminus \{u, v\}$.

Let us introduce the following specific notation, which allows us to express more than one dependence statement by a single term. Given a structure $\mathcal{P}$ over N, distinct variables $u, v \in N$ and a subset $U \subseteq N \setminus \{u, v\}$, the symbol $u \not\perp v | + U[\mathcal{P}]$ will be interpreted as the following:

$$u \not\perp v | + U[\mathcal{P}] \equiv \forall W \text{ such that } U \subseteq W \subseteq N \setminus \{u, v\} \text{ one has } u \not\perp v | W[\mathcal{P}].$$

In words, u and v are (conditionally) dependent in $\mathcal{P}$ given any superset of U. If U is empty, we write $*$ instead of $+\emptyset$.

$$u \not\perp v | *[\mathcal{P}] \equiv \forall W \text{ such that } W \subseteq N \setminus \{u, v\}\ u \not\perp v | W[\mathcal{P}].$$

Lemma 3.8. *Consider a structure $\mathcal{P}$. If for two distinct variables $u \leftrightarrow_{\mathcal{P}} v$, then $u \not\perp v | *[\mathcal{P}]$.*

Proof. Without affecting the generality, suppose $u \preceq_{\mathcal{P}} v$. Then by Lemma 3.6, $u \in K_{]v[}$. The sequence of markers $[K_{]v[}, u]$, $[K_{]v[}, v]$ is a W-avoiding trail for any $W \subseteq N \setminus \{u, v\}$. Hence $u \not\perp v | *[\mathcal{P}]$. □

As shown bellow, one can prove that the *connection set* is one of the characteristics common to all equivalent structures using this lemma. That is, $\mathcal{E}(\mathcal{P})$ is a characteristic property of all the structures from any equivalent class.

Lemma 3.9. *Let $\mathcal{P}$ be a structure over N. Then for any two distinct variables $u, v \in N$ such that $u \preceq_{\mathcal{P}} v$,*

$$u \notin S(K_{]v[}) \ \wedge\ u \perp v | S(K_{]v[})[\mathcal{P}] \Leftrightarrow u \not\leftrightarrow_{\mathcal{P}} v.$$

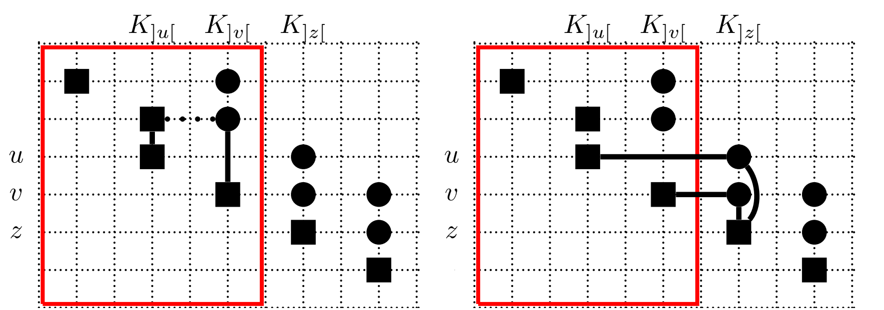
Proof. $\Rightarrow$ Suppose $u \perp\!\!\!\perp v | S(K_{]v[}) [\mathcal{P}]$ and $u \leftrightarrow_{\mathcal{P}} v$. This, however, contradicts Lemma 3.8, which asserts that $u \not\perp\!\!\!\perp v | * [\mathcal{P}]$ and therefore $u \not\perp\!\!\!\perp v | S(K_{]v[}) [\mathcal{P}]$ as well.

$\Leftarrow$ Suppose $u \not\leftrightarrow_{\mathcal{P}} v$. This excludes $u \in S(K_{]v[})$. Thus, assume $u \notin S(K_{]v[})$ and $u \not\perp\!\!\!\perp v | S(K_{]v[}) [\mathcal{P}]$. Since $u \leq_{\mathcal{P}} v$, and $S(K_{]v[}) \prec_{\mathcal{P}} v$, then according to Lemma 2.13 $u \perp\!\!\!\perp v | S(K_{]v[}) [\mathcal{P}[v]]$. By corollary 2.19 $u \perp\!\!\!\perp v | S(K_{]v[}) [\mathcal{P}]$, which contradicts the assumptions.

$\square$

Interestingly, notice that while a more general implication $u \perp\!\!\!\perp v | + S(K_{]v[}) [\mathcal{P}] \Rightarrow u \not\leftrightarrow v$ holds, the opposite one does not. One can find a counterexample of the opposite generalisation in the second part of the following example:

Example 3.10. *Let $\mathcal{P}$ be a structure with a persegram from Figure 9. Since $u \not\leftrightarrow_{\mathcal{P}} v$ then $u \perp\!\!\!\perp v | S(K_{]v[}) [\mathcal{P}]$ by Lemma 3.9. Let us check whether there is an $S(K_{]v[})$-avoiding trail in Figure 9a. We may restrict the searching area to an induced substructure $\mathcal{P}[\{u, v\} \cup S(K_{]v[})] = \mathcal{P}[v]$ by Corollary 2.19. The area corresponding to this substructure is highlighted. Since the only sequence of markers connecting u, v contains a horizontal connection corresponding to a variable from $S(K_{]v[})$, there is no $S(K_{]v[})$-avoiding trail in the persegram of $\mathcal{P}[v]$. Thus, $u \perp\!\!\!\perp v | S(K_{]v[}) [\mathcal{P}]$ by Corollary 2.19.*



(a) $u \perp\!\!\!\perp v | S(K_{]v[})$ (b) $u \not\perp\!\!\!\perp v | + S(K_{]v[})$

Figure 9: A contra-example that Lemma 3.9 cannot be generalised.

One can easily find an example that $u \not\leftrightarrow v \not\Rightarrow u \perp\!\!\!\perp v | + S(K_{]v[}) [\mathcal{P}]$ in Figure 9b. It is enough to realise that $\{z\} \cup S(K_{]v[})$ is just a special case of $+S(K_{]v[})$.

With the help of the previous lemma, one can prove the following important assertion.

Lemma 3.11. *Let $\mathcal{P}$ be a structure over N and $u, v \in N$ two distinct variables. Then*

$$u \leftrightarrow_{\mathcal{P}} v \Leftrightarrow u \not\perp\!\!\!\perp v | * [\mathcal{P}].$$

Proof. By Lemma 3.8, it will be enough to prove the implication ($\Leftarrow$). Suppose for contradiction that $u \not\perp\!\!\!\perp v | * [\mathcal{P}]$ and $u \not\leftrightarrow_{\mathcal{P}} v$; one can assume without loss of generality that $u \prec_{\mathcal{P}} v$. Then Lemma 3.9 leads to contradiction, since $u \perp\!\!\!\perp v | S(K_{]v[})$. $\square$

Corollary 3.12. *Let $\mathcal{P}, \mathcal{P}'$ be two structures over N. If $\mathcal{I}(\mathcal{P}) = \mathcal{I}(\mathcal{P}')$ then $\mathcal{E}(\mathcal{P}) = \mathcal{E}(\mathcal{P}')$.*

Remark 3.13. *Compositional model is, in fact, a multidimensional probability distribution and, as such, it can be represented by a Bayesian network as well. If one uses the conversion algorithm from the [1], then the structure of a created Bayesian network $G(N, E)$ - acyclic directed graph (dag) - induces the same independence model as the input compositional model structure $\mathcal{P}$. Moreover, the* connection *defined above corresponds precisely to the edge of the corresponding dag in case of the mentioned algorithm. I.e. $u \leftrightarrow_{\mathcal{P}} v \Leftrightarrow u \rightarrow v$ in G or $u \leftarrow v$ in G. This gives un a check that our conclusions are correct. Indeed, the set of connections $\mathcal{E}(\mathcal{P})$ (sometimes denoted as a* skeleton*) is a characteristic property of all dags equivalent with G by [3].*

Example 3.14. *In Example 3.3 the equivalence of different structures was discussed. The first two ($\mathcal{P}_1, \mathcal{P}'_1$) were equivalent, the second two ($\mathcal{P}_2, \mathcal{P}'_2$) were not. Let us look at that example again in the light of the previous corollary.*

1. *Let $\mathcal{P}_1, \mathcal{P}'_1$ be two simple structures depicted in Figure 6.*
   One may easily see that $\mathcal{E}(\mathcal{P}_1) = \mathcal{E}(\mathcal{P}'_1) = \emptyset$. The equality $\mathcal{I}(\mathcal{P}_1) = \mathcal{I}(\mathcal{P}'_1) = \{\langle u, v|\emptyset\rangle\}$ was shown in Example 3.3.

2. *On the other hand, consider structures $\mathcal{P}_2, \mathcal{P}'_2$ depicted in Figure 7. Notice that the corresponding connections are highlighted by arrows in Figures 8a and 8b. Due to Example 3.3 the reader knows that $\mathcal{I}(\mathcal{P}_2) \neq \mathcal{I}(\mathcal{P}'_2)$. Since $\mathcal{E}(\mathcal{P}'_2) = \mathcal{E}(\mathcal{P}_2) \cup \{\langle u, v\rangle\}$, the reason for non-equivalence is obvious now.*

3. *Consider structure $\mathcal{P}_2$ depicted in Figure 8a again. Is there any structure not equivalent with $\mathcal{P}_2$ but inducing the same connection set? Indeed, for an example see structure $\mathcal{P}_3$ depicted in Figure 8c. Observe that $u \not\perp\!\!\!\perp v|w[\mathcal{P}_2]$ but $u \perp\!\!\!\perp v|w[\mathcal{P}_3]$. Hence, $\mathcal{I}(\mathcal{P}_3) \neq \mathcal{I}(\mathcal{P}_2)$ while $\mathcal{E}(\mathcal{P}_3) = \mathcal{E}(\mathcal{P}_2)$.*

The 3rd part of Example 3.14 illustrates the fact that the same connection sets condition is necessary but not sufficient to guarantee the equivalence of respective structures. Therefore it is necessary to find an additional property invariant through a class of equivalent structures.

3.1.2. F condition set

We know that structures $\mathcal{P}_2, \mathcal{P}_3$ from the 3rd part of Example 3.14 are not equivalent despite the fact that $\mathcal{E}(\mathcal{P}_2) = \mathcal{E}(\mathcal{P}_3)$. Considering relation $\preceq_{\mathcal{P}}$, every structure induces a partial ordering of variables. One can easily verify that $u \prec_{\mathcal{P}_2} v \prec_{\mathcal{P}_2} w$ while $u \preceq_{\mathcal{P}_3} w \prec_{\mathcal{P}_3} v$. The induced variable ordering is different for non-equivalent structures. May the ordering of variables be some kind of characteristic property? Definitely not in this simple way: See Figure 6, where $\mathcal{I}(\mathcal{P}_1) = \mathcal{I}(\mathcal{P}'_1)$ while $u \prec_{\mathcal{P}_1} v$ and $u \succ_{\mathcal{P}'_1} v$.

It follows that two structures may induce different orderings of variables despite being equivalent. However, if we are only interested in the ordering of groups of specially connected variables, we obtain another property characteristic for a class of equivalent structures. This property is based on the so-called *F condition* defined below.

Definition 3.15. *Consider a structure $\mathcal{P}$ over N and three disjoint variables $u, v, w \in N$. We say that the triplet $\langle u, v|w\rangle$ is* F-condition *if*

$$\{u, v\} \prec_{\mathcal{P}} w, \{u, v\} \leftrightarrow_{\mathcal{P}} w, \text{ and } u \nleftrightarrow_{\mathcal{P}} v.$$

It is denoted by $u \eqslantless w \eqslantgtr v[\mathcal{P}]$. The set of triples $\mathcal{F}(\mathcal{P}) = \{\langle u, v|w\rangle : u \eqslantless w \eqslantgtr v[\mathcal{P}]\}$ is called F-condition set *induced by $\mathcal{P}$.*

The reason for calling the above-defined condition F-condition is very prosaic. Consider, for example, the structure $\mathcal{P}$ depicted in Figure 10. The reader can easily verify that $u \eqslantless w \eqslantgtr v[\mathcal{P}]$. Observe that w-avoiding trail connecting box-markers of u and v evokes a mirror image of letter F.

An example of F-condition can be found in $\mathcal{P}_2$ depicted in Figure 8a, where $u \eqslantless w \eqslantgtr v[\mathcal{P}_2]$. There is no F-condition in $\mathcal{P}_{2'}$ (Figure 8b) and $\mathcal{P}_3$ (Figure 8c).

Remark 3.16. *Lemma 3.6 says that conditions $u \prec_{\mathcal{P}} w$ and $u \leftrightarrow_{\mathcal{P}} w$ are equivalent to $u \in S(K^{\mathcal{P}}_{]w[})$. With regards to this, the previous definition may be reformulated in the following way:* Let $\mathcal{P}$ be a structure over N and $u, v, w \in N$. F-condition $u \eqslantless w \eqslantgtr v[\mathcal{P}]$ is a triplet of variables $\langle u, v|w\rangle$ such that $u, v \in S(K^{\mathcal{P}}_{]w[})$ and $u \nleftrightarrow_{\mathcal{P}} v$.

We have already shown that possessing the same *connection sets* is a necessary condition for equivalence of given structures. Therefore, when comparing two equivalent structures, the *connection set* may be considered as fixed. Now we show that the F-condition set is another characteristic property of a class of equivalent structures.

Lemma 3.17. *If three distinct variables $u, v, w \in N$ satisfy $\{u, v\} \leftrightarrow_{\mathcal{P}} w$ and $u \nleftrightarrow_{\mathcal{P}} v$ in a structure $\mathcal{P}$ over N, then*

$$u \eqslantless w \eqslantgtr v[\mathcal{P}] \Leftrightarrow u \not\perp\!\!\!\perp v| + w[\mathcal{P}].$$

Proof. $\Rightarrow$ Suppose $u \eqslantless w \eqslantgtr v[\mathcal{P}]$. Then $u, v \in S(K^{\mathcal{P}}_{]w[})$ by Remark 3.16. As one can see in Figure 10, the sequence of markers $[K_{]u[}, u]$, $[K_{]w[}, u]$, $[K_{]w[}, w]$, $[K_{]w[}, v]$, $[K_{]v[}, v]$ is a W-avoiding trail for every $W \subseteq N \setminus \{u, v\}$ such that $w \in W$. Hence, $u \not\perp\!\!\!\perp v|W[\mathcal{P}]$ for every such a W, which can be written as $u \not\perp\!\!\!\perp v| + w[\mathcal{P}]$.

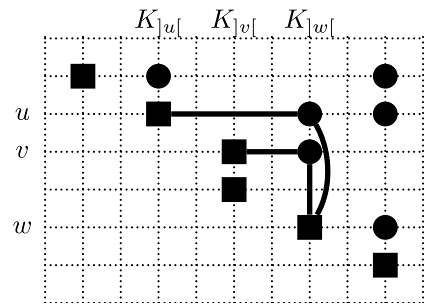


Figure 10: $u \not\perp v | + w$

$\Leftarrow$ To prove sufficiency by contradiction, assume that $u \not\perp v| + w[\mathcal{P}]$ and not $u \overset{\leftrightarrow}{\prec} w \overset{\leftrightarrow}{\succ} v[\mathcal{P}]$, which means that either $u \geq_{\mathcal{P}} w$ or $v \geq_{\mathcal{P}} w$. Assume without loss of generality $u \prec_{\mathcal{P}} v$ (equality may be omitted since $u \not\leftrightarrow_{\mathcal{P}} v$) and hence, $v \geq_{\mathcal{P}} w$. However, we may only consider the strict case $v \succ_{\mathcal{P}} w$. Indeed, otherwise $v \in R(K_{]w[})$ and $u \leftrightarrow_{\mathcal{P}} v$ which contradicts with the lemma assumption that $u \not\leftrightarrow_{\mathcal{P}} v$. The fact $v \succ_{\mathcal{P}} w$ implies that $w \in S(K^{\mathcal{P}}_{]v[})$ by assumption $w \leftrightarrow_{\mathcal{P}} v$ and Lemma 3.6. Since we assume that $u \not\leftrightarrow_{\mathcal{P}} v$, then $u \perp v | S(K_{]v[})\,[\mathcal{P}]$ by Lemma 3.9. This, however, contradicts with $u \not\perp v| + w[\mathcal{P}]$.

$\square$

Corollary 3.18. *Let $\mathcal{P}, \mathcal{P}'$ be two structures over N. If $\mathcal{I}(\mathcal{P}) = \mathcal{I}(\mathcal{P}')$ then $\mathcal{F}(\mathcal{P}) = \mathcal{F}(\mathcal{P}')$.*

Remark 3.19. *It has been mentioned in Remark 3.13 that there is an algorithm in [1] that enables us to create a dag G that induces the same independence model as a structure $\mathcal{P}$ - i.e. $\mathcal{I}(\mathcal{P}) = \mathcal{I}(G)$. Moreover, each edge in G corresponds to a connection from $\mathcal{P}$. Note that there is an edge $u \leftrightarrow v$ in G if and only if $u \to v$ or $u \leftarrow v$ in G. Since arrow orientation is given by relation $\prec_{\mathcal{P}}$ (if $u \prec_{\mathcal{P}} v$ and $u \leftrightarrow_{\mathcal{P}} v$ then $u \to v$ in G) in the conversion algorithm, then each F-condition defined above implies an* immorality *(vee-triple) in the respective dag G. Recall that ewe say that distinct nodes u, v, w form an immorality in a dag $G = (N, E)$ if $u \to w$ in G, $v \to w$ in G, and $u \not\leftrightarrow_G v$.*

We have derived two properties necessary for independence equivalence of given structures: *same connection and F-condition sets*. However, are these properties also sufficient to guarantee the equivalence of respective structures? Let us simply say that the answer is positive. However, since the goal of this paper is to present necessary conditions for equivalence of structures, we will not need this assertion here, and therefore we will not present its rather complex proof.

3.2. Column approach

One may disclose a possible non-equivalence of given structures with the help of characteristic properties introduced in the previous section. A problem arises when the considered structures are more complex and the rule of *same connection sets* is not so easily verifiable. It would be of special importance to have a rule concerning particular sets defining the structure instead of connections.

Is there such a condition? To cope with this question, we need the following definition.

Definition 3.20. *Let $\mathcal{P} = K_1, \ldots, K_n$ be a structure. A set K_i is a* non-trivial column *iff $R(K_i)$ is non-empty set. Otherwise it is a* trivial column *of $\mathcal{P}$. The symbol ntriv$(\mathcal{P})$ denotes the set of all non-trivial columns in $\mathcal{P}$. (ntriv$(\mathcal{P}) = \{K_i \in \mathcal{P} : R(K_i) \neq \emptyset\}$)*

Definition 3.21. *Let $\mathcal{P} = K_1, \ldots, K_n$ be a structure and $U \subseteq (K_1 \cup \ldots \cup K_n)$. The set U is* non-trivial in *$\mathcal{P}$ iff there is a $K_i \in \mathcal{P}$ such that $U \subseteq K_i$ and $R(K_i) \cap U \neq \emptyset$. Otherwise the set U is* trivial *in $\mathcal{P}$.*

Remark 3.22. *Observe that, since $R(K^{\mathcal{P}}_i) \neq \emptyset$ for nontrivial $K^{\mathcal{P}}_i$, it is obvious that ntriv$(\mathcal{P}) \leq |N|$.*

The following lemma proves that the set of mutually connected variables takes an important role in the structure. It is a consequence of Lemma 3.6.

Lemma 3.23. *Let U be a non-empty set of mutually connected variables in $\mathcal{P}$ ($u \leftrightarrow_{\mathcal{P}} u'$ for all $u, u' \in U$). Then U is a non-trivial set in $\mathcal{P}$.*

Proof. Choose $u \in U$ such that $u \geq_{\mathcal{P}} u'$ for all other $u' \in U$. This choice is always possible and ensures that $U \subseteq K^{\mathcal{P}}_{]u[}$. Indeed, $u \leftrightarrow_{\mathcal{P}} u'$ by assumption and therefore $u' \in K^{\mathcal{P}}_{]u[}$ by Lemma 3.6 for all $u' \in U \setminus \{u\}$. Since $u \in R(K_{]u[})$ by definition of function $] \cdot [_{\mathcal{P}}$, then U is a non-trivial set in $\mathcal{P}$ by Definition 3.21. $\square$

With the help of the previous lemma, one can prove the following interesting assertion concerning non-trivial sets and class of independence equivalent structures.

Lemma 3.24. *If a set of variables U is non-trivial in a structure $\mathcal{P}$, then it is non-trivial in every structure equivalent with $\mathcal{P}$.*

Proof. Assume that $\mathcal{P}'$ and $\mathcal{P}$ are equivalent. Then $\mathcal{E}(\mathcal{P}) = \mathcal{E}(\mathcal{P}')$ and $\mathcal{F}(\mathcal{P}) = \mathcal{F}(\mathcal{P}')$ according to Corollaries 3.12 and 3.18. The non-triviality of U implies the existence of a column $K^{\mathcal{P}}_{]w[}$ such that $w \in U$ and $U \subseteq K^{\mathcal{P}}_{]w[}$. Then $u \preceq_{\mathcal{P}} w$ and $u \leftrightarrow_{\mathcal{P}} w$ for all $u \in (U \setminus \{w\})$ by Lemma 3.6. Let $M \subseteq U$ be a maximal subset of mutually connected variables in $\mathcal{P}$ such that both $R(K^{\mathcal{P}}_{]w[}) \cap U \subseteq M$ and $M \leftrightarrow_{\mathcal{P}} u$ for all $u \in U \setminus M$. Put $V = U \setminus M$. Observe that not only $M \neq \emptyset$ ($w \in M$) but also $V \prec_{\mathcal{P}} w$. Indeed, suppose that $\exists v \in V$ such that $v \geq_{\mathcal{P}} w$. Then $v \in R(K^{\mathcal{P}}_{]w[})$ by definition of w, which contradicts the choice of M. One can distinguish two cases: $V = \emptyset$ and $V \neq \emptyset$.

If $V = \emptyset$ then $U = M$ is a set of mutually connected variables in $\mathcal{P}'$ by $\mathcal{E}(\mathcal{P}) = \mathcal{E}(\mathcal{P}')$. Therefore U is non-trivial in $\mathcal{P}'$ by Lemma 3.23.

Suppose now $V \neq \emptyset$. Since M is a set of mutually connected variables in $\mathcal{P}'$ by $\mathcal{E}(\mathcal{P}) = \mathcal{E}(\mathcal{P}')$, then M is non-trivial by Lemma 3.23 in $\mathcal{P}'$ and therefore $\exists m \in M$ such that $M \subseteq K^{\mathcal{P}'}_{]m[}$. The next step is to prove that $V \subset K^{\mathcal{P}'}_{]m[}$ as well.

Assume for a contradiction that $\exists v \in V$ such that $v \notin K^{\mathcal{P}'}_{]m[}$. There exists $v' \in V \setminus \{v\}$ such that $v' \not\leftrightarrow_{\mathcal{P}} v$ (otherwise $v \in M$). Considering the fact $V \prec_{\mathcal{P}} w$ and $\{v, v'\} \leftrightarrow_{\mathcal{P}} w$, we get $v \lnsim w \gnsim v'[\mathcal{P}]$ and $v \lnsim w \gnsim v'[\mathcal{P}']$ by $\mathcal{F}(\mathcal{P}) = \mathcal{F}(\mathcal{P}')$. The fact that $w \in M \subseteq K^{\mathcal{P}'}_{]m[}$ implies $w \preceq_{\mathcal{P}'} m$ by Lemma 3.6. Together with $v \prec_{\mathcal{P}'} w$ (because of $v \lnsim w \gnsim v'[\mathcal{P}']$), it follows that $v \prec_{\mathcal{P}'} m$. Moreover, $v \leftrightarrow_{\mathcal{P}'} m$ by definition of M and Corollary 3.12, and therefore $v \in K^{\mathcal{P}'}_{]m[}$ by Lemma 3.6, which contradicts the assumption.

Hence, $V \subset K^{\mathcal{P}'}_{]m[}$. Thus $U = V \cup M \subseteq K^{\mathcal{P}'}_{]m[}$ and $m \in R(K^{\mathcal{P}'}_{]m[})$ which guarantees the non-triviality of U in $\mathcal{P}'$ by Definition 3.21. $\square$

Observe that there is a close relationship between non-trivial columns and non-trivial sets of variables. In fact, an arbitrary non-trivial column $K^{\mathcal{P}}$ is a non-trivial set $U = K^{\mathcal{P}}$ as well.

Lemma 3.25. *Having fixed structure $\mathcal{P}$, the maximal non-trivial sets (with respect to inclusion) in $\mathcal{P}$ coincide with maximal sets in $ntriv(\mathcal{P})$ (with respect to inclusion), that is maximal columns with at least one box-marker.*

Proof. To prove this lemma it is enough to realise that every non-trivial column $K_i \in ntriv(\mathcal{P})$ represents a non-trivial set of variables $U = K_i$ at the same time. Similarly, an existence of some non-trivial set U implies the existence of a non-trivial column $K^{\mathcal{P}}$ such that $U \subseteq K^{\mathcal{P}}$ by Definition 3.21.

Suppose for a contradiction the existence of a maximal non-trivial column K (which coincides with a non-trivial set U) and some non-trivial set V such that $U \subset V$. The non-triviality of V implies the existence of $K' \in ntriv(\mathcal{P})$ such that $V \subseteq K'$. Then $K \subset K'$ which contradicts the fact that K is maximal non-trivial column with respect to inclusion. In particular, every maximal non-trivial column is a maximal non-trivial set.

Similarly consider a maximal non-trivial set U. Since U is non-trivial then there exists a non-trivial column K_i where $U \subseteq K_i$ by assumption. There exists a maximal non-trivial column K_j with $K_i \subseteq K_j$ (possibly $i = j$). Thus, $U \subseteq K_j$. As K_j is a non-trivial set, necessarily $U = K_j$ for otherwise U is not maximal (non-trivial set). Thus U coincides with a maximal column. $\square$

Definition 3.26. *For a structure $\mathcal{P}$, its* strong core $C^*(\mathcal{P})$ *is the set of maximal non-trivial columns with respect to inclusion. ($C^*(\mathcal{P}) = \{K_i \in ntriv(\mathcal{P}) : \nexists K_j \in ntriv(\mathcal{P})$ such that $K_i \subset K_j\}$.)*

Observe that, for a given structure $\mathcal{P}$, its strong core $C^*(\mathcal{P})$ does not contain any trivial columns from $\mathcal{P}$. $C^*(\mathcal{P}) \subseteq ntriv(\mathcal{P})$

Corollary 3.27. *Let $\mathcal{P}$ be a structure over N. Then $C^*(\mathcal{P}) = C^*(\mathcal{P}')$ for every equivalent structure $\mathcal{P}'$.*

Proof. Since by Lemma 3.24 non-trivial sets are same for equivalent structures, the classes of maximal non-trivial sets in them coincide. Thus maximal non-trivial columns (by Lemma 3.25) in them coincide. □

Remark 3.28. *Unlike the previously discussed invariants in Remarks 3.13 and 3.19, this characteristic property does not correspond to any standard characteristics of equivalent dags. Still, given the definition,* strong structure core *could correspond to a set of maximal families induced by corresponding $G = (N, E)$. Note that by family $fam(u)$ we understand the set $u \cup pa(u)$ where $pa(u) = \{v \in N : (v \to u) \in E\}$.*

Recall that one can generate all dags which are equivalent to a given one with the help of the so called legal arrow reversal. *By a* legal arrow reversal *we understand the change of dag G into dag G' by replacement of an arrow $u \to v$ (in G) by $u \leftarrow v$ (in G') under the condition that $pa_G(u) \cup u = pa_G(v)$. If $fam_G(v) = V = pa_G(u) \cup \{u, v\}$ belongs to maximal families (with respect to inclusion) then it belongs to maximal families in G' as well. Indeed, since $pa_{G'}(u) = pa_G(u) \cup v$ then $fam_{G'}(u) = V$. Since no other arrow changes, then all other families remain the same and then V belongs to maximal families in G' and in all other equivalent dags.*

Remark 3.29. *Based on work with a variety of equivalent structures, it appears that the* strong core *definition could be modified. The new definition of the core would then be as follows: For a structure $\mathcal{P}$, its* weak core *$C(\mathcal{P})$ is a set of $K_i \in ntriv(\mathcal{P})$ such that $K_i \neq S(K_j)$ for every $K_j \in ntriv(\mathcal{P})$.*

Note that the extended weak core also includes not only strong core but also all the columns that are not sharp subsets of another non-trivial column. However, we are not able to give a simple clear proof that the weak core is another characteristic of structure equivalence.

4. Conclusion

This paper started with a brief introduction on how to read unconditional independencies induced by a structure of a compositional model. We then introduced, step by step, several properties necessary for independence equivalence of the relevant structures of compositional models: Two structures, if equivalent, must have the same *connection sets* and *F-condition sets*. Since it is difficult to verify the existence of the same *connection set* in more complex structures, we apply a new approach based on structure columns.

Based on this, we understand a column with a box-marker as a *non-trivial* set of variables and we have shown that every non-trivial set in one structure has to be non-trivial in all structures equivalent with it. Columns that have to be in all persegrams of equivalent structures are called *structure core*. This offers a powerful tool for determining the possible non-equivalence of given structures.

Acknowledgements

This work was supported by National Science Foundation of the Czech Republic under Grants No. ICC/08/E010, and 201/09/1891, and by Ministry of Education under Grants No. 1M0572 and 2C06019.

Let me thank Radim Jiroušek for his cooperation in solving various problems, for reading the manuscript in its various stages and numerous ideas that improved the paper. Also let me thank to Milan Studený for his careful reading the manuscript and for helpful and valuable suggestions which led to a simplification in the last part of the paper.

References

[1] R. Jiroušek. *Multidimensional Compositional Models. Part 1: Introduction*. Internal Publication 2006/4 ÚTIA AV ČR, Prague (2006)

[2] R. Jiroušek. *Persegrams of Compositional Models Revisited: conditional independence*. Proceedings of the 12th International Conference on Information Processing and Management of Uncertainty in Knowledge-based Systems, Malaga, (2008).

[3] T. Kočka, R. R. Bouckaert, M. Studený. *On the Inclusion Problem*. Research report 2010, ÚTIA AV ČR, Prague (2001).

[4] V. Bína, R. Jiroušek. *Marginalization in Multidimensional Compositional Models*, Kybernetika vol.42, 4 (2006), p. 405-422

[5] V. Kratochvíl. *Equivalence Problem in Compositional Models*. Proceedings of the 8th Workshop on Uncertainty Processing, Eds: Kroupa T., Vejnarová J., Prague, (2009).

[6] V. Kratochvíl. *Motivation for different characterization of Equivalent Persegrams*. Proceedings ot the 12th Czech-Japan seminar on Data Analysis and Decision Making under Uncertainty, Eds: Novák V., Pavliska V., Štěpnička M., Ostrava, (2009).

[7] V. Kratochvíl. *Different Approaches of Study Direct Equivalence Characterization*, Doktorandské dny 2009, Eds: Ambroř P., Masáková Z., Prague (2009).