Motivation for different characterizations of Equivalent Persegrams

Václav Kratochvíl*
Institute of Information Theory and Automation
Academy of Sciences of the Czech Republic
velorex@utia.cz

Abstract

Structure of each Compositional model can be visualized by a tool called persegram. Every persegram over a finite non-empty set of variables N induces an independence model over N, which is a list of conditional independence statements over N. The *equivalence problem* is how to characterize (in graphical terms) whether all independence statements in the model induced by persegram $\mathcal{P}$ are in the model induced by a second persegram $\mathcal{P}'$ and vice versa. In the previous paper [6] some kind of direct characterization of equivalence was done. We introduced two invariant properties of equivalent persegrams. Are these invariants sufficient to decide of equivalence of given persegrams? This question has not been answered yet.

In this paper we give the motivation and introduction for indirect characterization of equivalence. We have found three operations on persegram remaining induced independence model. By combining them together, one can generate a class of equivalent models. We are not sure whether one can generate the *whole* class. This problem is closely connected with the above mentioned problem of invariant 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 are an alternative approach to Graphical Markov Models. These models

*The research was partially supported by Ministry of Education of the Czech Republic under grant no. 2C06019, 1M0572, and by Czech Science Foundation under grants no. ICC/08/E010 and 201/09/1891.

are generated by a sequence (generating sequence) of low-dimensional distributions, which, composed together, create a distribution - the so called *Compositional model*. Moreover, while a model is composed together, a system of (un)conditional independencies is simultaneously introduced by the structure of the generating sequence.

The structure can be visualized by a tool called *persegram* and one can read induced independencies directly using this tool. That is why we can say that every persegram over a finite non-empty set of variables N induces an *independence model* over N - a list of conditional independence statements over N. The *equivalence problem* is how to characterize (in graphical terms) whether all independence statements in the model induced by persegram $\mathcal{P}$ are also in the independence model induced by a second persegram $\mathcal{P}'$ and vice versa.

2 Notation and Basic Properties

Throughout the paper the symbol N will denote a non-empty set of finite-valued *variables*. From the next chapter on, variables will be represented by markers of a persegram. All probability distributions of this variables will be denoted by Greek letters (usually π, κ); thus for $K \subset N$, we consider a distribution (a probability measure over K) $\pi(K)$ which is defined for variables K. When several distributions will be considered, we shall distinguish them by indices. For a probability distribution $\pi(K)$ and $U \subset K$ we will consider a *marginal distribution* $\pi(U)$.

The following conventions will be used throughout the paper. Given sets $K, L \subset N$ the juxtaposition KL will denote their union $K \cup L$. The following symbols will be reserved for special subsets of N: K, R, S. The symbol U, V, W, Z will be used for general subsets of N. The symbol $|U|$ will be used to denote the number of elements of a finite set U, that is, its *cardinality*. u, v, w, z denotes variables as well as singletons $\{x\}, \ldots$

Independence and dependence statements over N correspond to special *disjoint triples* over N. The symbol $\langle U, V|Z\rangle$ denotes a triplet of pairwise disjoint subsets U, V, Z of N. This notations anticipates the intended meaning: the set of variables U is conditionally independent or dependent of the set of variables V given the set of variables Z. This is why the third set Z is separated by a straight line: it has a special meaning of the conditioning set. The symbol $\mathcal{T}(N)$ will denote the class of all disjoint triplets over N:

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

To describe how to compose low-dimensional distributions to get a distribution of a higher dimension we use the following operator of composition.

Definition 2.1. For arbitrary two distributions $\pi(K)$ and $\kappa(L)$ their *composition* is given by the formula

$$\pi(K) \triangleright \kappa(L) = \begin{cases} \frac{\pi(K)\kappa(L)}{\kappa(K \cap L)} & \text{if } \pi^{\downarrow K \cap L} \ll \kappa^{\downarrow K \cap L}, \\ \text{undefined} & \text{otherwise,} \end{cases}$$

where the symbol $\pi(M) \ll \kappa(M)$ denotes that $\pi(M)$ is *dominated* by $\kappa(M)$, which means (in the considered finite setting)

$$\forall x \in \times_{j \in M} \mathbf{X}_j; (\kappa(x) = 0 \Longrightarrow \pi(x) = 0).$$

The result of the composition (if defined) is a new distribution. We can iteratively repeat the process of composition to obtain a multidimensional distribution - a model approximating the original distribution with corresponding marginals. That is why the multidimensional distribution (and the whole theory as well) is called *Compositional model*. To describe such a model it is sufficient to introduce an ordered system of low-dimensional distributions $\pi_1, \pi_2, \ldots, \pi_n$. If all compositions are defined, we call this ordered system a *generating sequence*.

From now on, we consider generating sequence $\pi_1(K_1), \pi_2(K_2), \ldots, \pi_n(K_n)$ which defines a distribution (where the operator $\triangleright$ is applied from left to right)

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

Therefore, whenever distribution π_i is used, we assume it is defined for variables K_i. In addition, each set K_i can be divided into two disjoint parts. We denote them R_i and S_i with the following sense:

$$R_i = K_i \backslash (K_1 \cup \ldots \cup K_{i-1}), S_i = K_i \cap (K_1 \cup \ldots \cup K_{i-1})$$

.

R_i denotes variables from K_i with the first appeared with respect to the sequence (meaning from left to right). S_i denotes the already used.

2.1 Graphical concepts

It is well-known that one can read conditional independence relations of a Bayesian network from its graph. A similar technique is used in compositional models. An appropriate tool for this is a *persegram*. Persegram is used to visualize the structure of a compositional model and is defined bellow.

Definition 2.2. Persegram $\mathcal{P}$ *of a generating sequence is a table in which rows correspond to variables (in an arbitrary order) and columns to low-dimensional distributions; ordering of the columns corresponds to the generating sequence ordering. A position in the table is marked if the respective distribution is defined for the corresponding variable. Markers for the first occurrence of each variable (i.e., the leftmost markers in rows) are squares (we call them* box-markers*) and for other occurrences there are* bullets.

Since the markers in the i-th column represent variables K_i, we denote markers in i-th column as K_i. Box-markers in i-th column of $\mathcal{P}$ are denoted like R_i and bullets like S_i. $K_i = R_i \cup S_i$. This notation is purposely in accordance with notation of variable sets in generating sequences to simplify readability and lucidity of the text.

Persegrams are usually denoted by $\mathcal{P}$ and if it is not specified otherwise, $\mathcal{P}$ corresponds to the generating sequence $\pi_1(K_1),\ldots,\pi_n(K_n)$ where $K_1\cup\ldots\cup K_n = N$. We say that $\mathcal{P}$ is *defined over* N. (i.e. $\mathcal{P}$ over N has n columns with markers $K_1,\ldots,K_n$ where $K_1\cup\ldots\cup K_n = N$.)

To simplify the notation we will use the following symbol: Let $\mathcal{P}$ be a persegram over N. We introduce a function $][_{\mathcal{P}}\colon N\to\mathbb{N}$, which for every variable $u\in N$ returns the index of set K_i with the first appearance of u in the persegram $\mathcal{P}$. Due to the previously established notation can be said that $K_{]u[_{\mathcal{P}}}$ is a column K_i where $u\in R_i$. In other words: $]u[_{\mathcal{P}}= i : u\in R_i$.

Definition 2.3. *Let $\mathcal{P}$ be a persegram over N and $\preceq_{\mathcal{P}}$ a binary relation. For arbitrary $u,v\in N$ we denote $u\preceq_{\mathcal{P}} v$ if $]u[_{\mathcal{P}}\leq]v[_{\mathcal{P}}$. Moreover we introduce the relation $\prec_{\mathcal{P}}$: $u\prec_{\mathcal{P}} v\Leftrightarrow u\preceq_{\mathcal{P}} v$ AND $v\not\preceq_{\mathcal{P}} u$.*

The following convention will be used throughout the paper: Given variables $u,v,w\in N$ and $\mathcal{P}$ over N, the term $u,v\prec_{\mathcal{P}} w$ denotes that $u\prec_{\mathcal{P}} w$ and $v\prec_{\mathcal{P}} w$. The symbol $\mathcal{P}$ may be omitted, if the content is clear.

2.2 Conditional independence

Conditional independence statements over N induced by the structure of Compositional model can be read from its persegram. Such independence is indicated by the absence of a *trail connecting or avoiding relevant markers*. It is defined below.

Definition 2.4. *Consider a persegram over N and a subset $Z\subset N$. A sequence of markers $m_0,\ldots,m_t$ is called a* Z-avoiding trail *that connects m_0 and m_t if it meets the following 4 conditions:*

1. *for each $s=1,\ldots,t$ a couple (m_{s-1},m_s) is 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 (one of the markers is 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: 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 also say that these variables are connected by a Z-avoiding trail. *Suppose $\langle U,V|Z\rangle\in\mathcal{T}(N)$ is a disjoint triplet over N. One says that U and V are conditionally dependent by Z, written $U\not\perp\!\!\!\perp V|Z[\mathcal{P}]$, if there exists a Z-avoiding trail between variable $u\in U$ and variable $v\in V$ in $\mathcal{P}$. In the opposite case one says that U and V are conditionally independent by Z in $\mathcal{P}$, written $U\perp\!\!\!\perp V|Z[\mathcal{P}]$.*

We also say that $\langle U, V|Z\rangle$ *is* represented *in* $\mathcal{P}$. *The induced independence model* $\mathcal{I}(\mathcal{P})$ *and the induced dependence model* $\mathcal{D}(\mathcal{P})$ *are defined as follows:*

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

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

Example 2.5. *Consider persegram from Figures 1 and 2.*

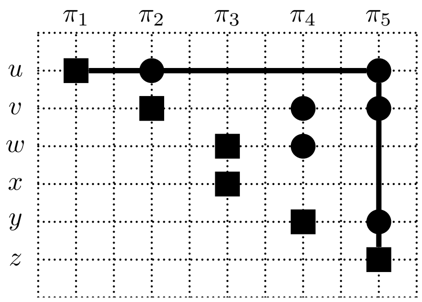

Figure 1: $\mathcal{P} : u \not\perp\!\!\!\perp z|\emptyset, u \not\perp\!\!\!\perp z|v$

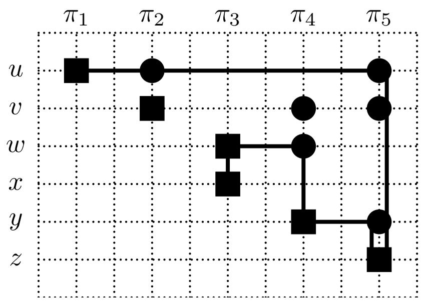

Figure 2: $\mathcal{P} : u \not\perp\!\!\!\perp x|z$

In Figure 1 a $\emptyset$*-avoiding trail is depicted. Therefore* $u \not\perp\!\!\!\perp z|\emptyset$. *Moreover, one can replace* $\emptyset$ *by any subset of* $\{v, w, x, y\}$ *which is avoiding* Z *as well. In Figure 2, there is depicted another trail connecting* u *and* x. *Therefore* $u \not\perp\!\!\!\perp x|z$. *On the contrary to Figure 1, one can not replace* z *by any other variable except* v. *Otherwise, the condition 3. from the Definition 2.4 will be corrupted. (i.e.* $u \perp\!\!\!\perp x|y[\mathcal{P}]$ *for example)*

The following specific notation for certain composite dependence statements will be useful. Given a persegram $\mathcal{P}$ over N, distinct variables $u, v \in N$ and disjoint set $U \subseteq N \setminus \{u, v\}$ the symbol $u \not\perp\!\!\!\perp v| + U[\mathcal{P}]$ will be interpreted as the condition

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

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

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

We give a certain graphical characterization of composite dependence statements of this kind below.

3 Equivalence problem

By the equivalence problem we understand the problem how to recognize whether two given persegrams $\mathcal{P}_1, \mathcal{P}_2$ over N induce the same independence model ($\mathcal{I}_{\mathcal{P}_1} =$

$\mathcal{I}_{\mathcal{P}_2}$). It is of special importance to have an easy rule to recognize that two persegrams are equivalent in this sense and an easy way to convert $\mathcal{P}_1$ into $\mathcal{P}_2$ in terms of some elementary operations on persegrams. Another very important aspect is the ability to generate all persegrams which are equivalent to a given persegram.

Definition 3.1. *Persegrams $\mathcal{P}_1, \mathcal{P}_2$ (over the same variable set N) are called* independence equivalent*, if they induce the same independence model $\mathcal{I}_{\mathcal{P}_1} = \mathcal{I}_{\mathcal{P}_2}$.*

Remark 3.2. *One may easily see that the above mentioned definition could be formulated with the term of dependence model.* Persegrams $\mathcal{P}_1, \mathcal{P}_2$ (over the same variable set N) are independence equivalent, iff $\mathcal{D}_{\mathcal{P}_1} = \mathcal{D}_{\mathcal{P}_2}$. *This alternative is used in most proofs primarily.*

3.1 Direct characterization

The solution of equivalence problem can be done in several ways. Some kind of *direct characterization* of equivalence follows was done in the paper [6] where we introduced two invariant properties of equivalent persegrams. Let us remind these invariant together with necessary definitions of *connection* and *ordering condition*. Proofs can be found in [6].

Definition 3.3. *Let $\mathcal{P}$ be a persegram over N and $u, v \in N$ be two distinct variables, and $u \preceq_{\mathcal{P}} v$. u, v are* connected *in $\mathcal{P}$ ($u \leftrightarrow v[\mathcal{P}]$) if $u \in K_{]v[}$. The set of all pairs $\mathcal{E}(\mathcal{P}) = \{\langle u, v\rangle : u, v \in N, u \leftrightarrow v[\mathcal{P}]\}$ is called a* connection set *of $\mathcal{P}$.*

Lemma 3.4. *Let $\mathcal{P}$ be a persegram over N and $u, v \in N$ are distinct variables. Then*

$$u \leftrightarrow v[\mathcal{P}] \Leftrightarrow u \not\perp\!\!\!\perp v \mid *\, [\mathcal{P}].$$

Definition 3.5. *Let $\mathcal{P}$ be a persegram over N. An* Ordering condition *is a triplet of variables $u, v, w \in N$ where $u, v \prec w$; $u, v \leftrightarrow w$; and $u \not\leftrightarrow v$ in $\mathcal{P}$. Such an ordering condition is denoted by $[u, v] \prec w[\mathcal{P}]$.*

Lemma 3.6. *Let $\mathcal{P}$ be a persegram over N, $u, v, w \in N$ distinct nodes. Then*

$$[u, v] \prec w \Leftrightarrow u \not\perp\!\!\!\perp v \mid +\, w[\mathcal{P}].$$

The previous lemmata show two invariant properties of equivalent persegrams. Two persegrams, if equivalent, have the same set of connections and induce the same set of ordering conditions.

Corollary 3.7. *Let $\mathcal{P}, \mathcal{P}'$ be two persegrams over N. If $\mathcal{I}_{\mathcal{P}} = \mathcal{I}_{\mathcal{P}'}$ then $\mathcal{E}(\mathcal{P}) = \mathcal{E}(\mathcal{P}')$ and they induce the same set of ordering conditions.*

The question is: Does this implication hold also in the opposite direction? I.e. if two persegrams $\mathcal{P}, \mathcal{P}'$ over the same set N induce the same ordering conditions and $\mathcal{E}(\mathcal{P}) = \mathcal{E}(\mathcal{P}')$, are $\mathcal{P}, \mathcal{P}'$ independence equivalent? The answer for this question is still unknown. Despite the fact that all experiments confirm this theory, the formal proof has not been finished yet.

We can introduce a set of operations on persegram preserving induced independence model with the help of these properties.

3.2 IE operations

Suppose we are interested in the solution of the third part of the Equivalence problem. We want to find some operations on persegram that preserve independence model. What are the possible operations on persegram?

If you realize the Definition 2.2 *Change row ordering* has no effect on the persegram. By adding a row one adds a new variable into a model. It makes sense to consider equivalent persegrams over the same set of variables N only and therefore *Change row amount* does not need to be consider.

Add/remove marker: Suppose one adds a bullet corresponding to $v \in N$ into $K_{]u[}$. Then by Definition 3.3 one adds a connection $u \leftrightarrow v$ that was not there before. By adding of a box-marker corresponding to v into $K_{]u[}$ new connections $v \leftrightarrow K_{]u[} \setminus S_{]v[}$ appear and $v \leftrightarrow S_{]v[} \setminus K_{]u[}$ disappear simultaneously. It will not happen if $K_{]u[} = S_{]v[}$ (see Example 3.8).

Change column ordering could be possible under special circumstances as well as *Change column amount.* One can always add a column consisting of bullets only.

Example 3.8. *Let $\mathcal{P}_1$ be a persegram over N, $u, v \in N$, and $K_{]u[} = S_{]v[}$. These two columns are given in Figure 3. Then one can add a box-marker corresponding to v into $K_{]u[}$ to get persegram $\mathcal{P}_2$. Since one adds a new box-marker, the old one turns into bullet and $v \leftrightarrow S_{]v[} \setminus K_{]u[}$ disappear. $v \leftrightarrow K_{]u[} \setminus S_{]v[}$ appears in $K_{]u[}$. These connections are highlighted in Figure 3. Since $K_{]u[} = S_{]v[}$ then $\mathcal{E}(\mathcal{P}_1) = \mathcal{E}(\mathcal{P}_2)$.*

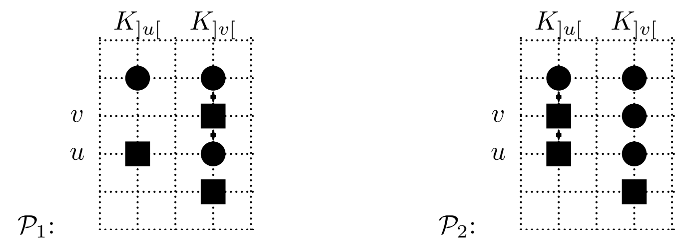

Figure 3: Since $K_{]u[} = S_{]v[}$, one can add box-marker of v into $K_{]u[}$.

The previous discussion shows that we can restrict to the operations with columns only. Moreover, a box marker can be added/removed under special circumstances.

Four different simple operations on persegram preserving independence model were discovered. We call them *IE operations* (Independence Equivalent). These operations can be divided into two groups according to how they do by columns:

Either changing ordering (this group is called *permutations*) or adding/removing them (*extensions/reductions*).

The idea of proofs is very simple. We will prove that every Z-avoiding trail either remains the same or it can be easily transformed according to the Remark 3.2. We consider the following operations:

Definition 3.9. *Let $\mathcal{P}$ be a persegram over N and two adjacent columns K_i, K_{i+1}. The so called* IE operations *are the following set of operations with columns.*

- Independent permutation: *We can swap two columns K_i, K_{i+1}, if no box-marker turns into bullet and vice-versa. ($\cup_{j=1}^{i-1} Kj \supseteq K_i \cap K_{i+1}$)*
- Intersection permutation: *We can swap two columns K_i, K_{i+1}, if all bullets $S_i \cup S_{i+1}$ belong to its intersection. ($S_i \cup S_{i+1} \subseteq K_i \cap K_{i+1}$)*
- *Removing of a column containing bullets only is called* Bullets extension/reduction. *($K_i = S_i$)*
- *Removing of a column i, which is a subset of the column $i+1$ that has box-markers elsewhere only, is called* Subset extension/reduction. *($K_i = S_{i+1}$.)*

Claim that Independent permutation preserve the model was demonstrated in the [5]. It is a little bit more complicated in case of *Intersection permutation.*

Lemma 3.10. *Let $\mathcal{P}$ be a persegram over N. If $\mathcal{P}'$ arises from $\mathcal{P}$ by applying of* Intersection permutation *then $\mathcal{I}_{\mathcal{P}} = \mathcal{I}_{\mathcal{P}'}$.*

Proof. We show that $\mathcal{D}_{\mathcal{P}} = \mathcal{D}_{\mathcal{P}'}$. By the definition of *Intersection Permutation* the role of $\mathcal{P}$ and $\mathcal{P}'$ is interchangeable. It suffices to verify $\mathcal{D}_{\mathcal{P}} \subseteq \mathcal{D}_{\mathcal{P}'}$. Suppose, that $\mathcal{P}'$ arises from $\mathcal{P}$ by applying of *Intersection permutation* on two adjacent columns K_i, K_{i+1}.

Suppose there is a Z-avoiding trail $u \not\perp\!\!\!\perp v|Z[\mathcal{P}]$ Without loss of generality, consider a Z-avoiding trail τ which involves the minimal number of markers among trails of this type. Since $\mathcal{P}$ differs from $\mathcal{P}'$ in K_i, K_{i+1} columns only (its order and markers form), we show how to convert that part of τ into $\mathcal{P}'$ and therefore that $u \not\perp\!\!\!\perp v|Z[\mathcal{P}']$.

1. Red-line the sequence of markers τ in $\mathcal{P}$.
2. Apply the *Intersection permutation* on K_i, K_{i+1} columns of $\mathcal{P}$ ($K_i[\mathcal{P}] \Rightarrow K'_{i+1}[\mathcal{P}']$, $K_{i+1}[\mathcal{P}] \Rightarrow K'_i[\mathcal{P}']$) and move the corresponding parts of τ together with columns K_i, K_{i+1} to create τ'.
3. If exists a vertical connection of τ' not adjacent to any box-marker (in K'_{i+1}) move it into K'_i column. (It is possible since $(S'_i \cup S'_{i+1} = S_{i+1} \cup S_i) \subseteq (K_i \cap K_{i+1} = K'_{i+1} \cap K'_i)$)
4. Reduce or extend the corresponding horizontal connection.

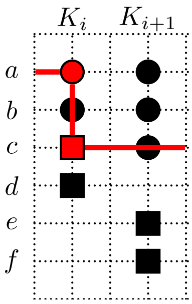

Figure 4: Step 1

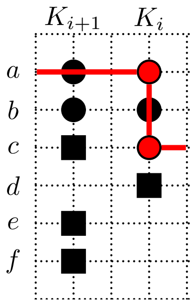

Figure 5: Step 2

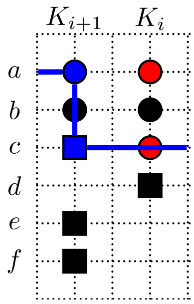

Figure 6: Steps 3,4

The execution phase of the algorithm is demonstrated on figures 4, 5, and 6. The only part where τ' differs from τ (and where it can breaks the conditions of definition 2.4) is in K_i, K_{i+1}. Since τ satisfied all conditions of Definition 2.4 then one can easily see that, except the condition 4, τ' does as well - *Horizontal and vertical connections regularly alternate.* Let us test this condition too.

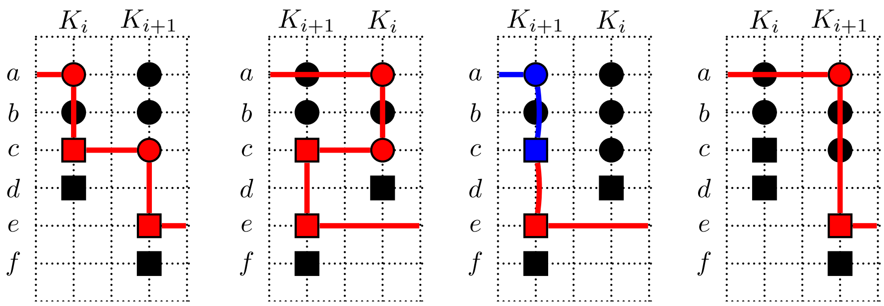

Figure 7: First trail, which leads to two verticals, contradicts with the fact that we use the shortest trail. The fourth trail which is shorter.

Suppose that vertical and horizontal connections do not regularly alternate in τ'. It may happen only if in step 3, after moving the two vertical connections appear in the same column. However, it can not happen because in that case τ is not the shortest sequence avoiding Z (see on figure 7), which contradicts with assumptions. $\square$

Lemma 3.11. *Let $\mathcal{P}$ be a persegram. If $\mathcal{P}'$ arises from $\mathcal{P}$ by applying of Subset extension/reduction then $\mathcal{I}_{\mathcal{P}} = \mathcal{I}_{\mathcal{P}'}$.*

Proof. This lemma can by proved the same way as lemma 3.10, or one can realize that Subset extension/reduction can be spread out into Intersection permutation and Bullets extension/reduction, where both of them preserve Independence model $\mathcal{M}_P$. $\square$

We have found four operations on persegrams preserving the induce independence model till now. Two of them change the columns ordering, the other two extend/reduce its length. With the help the extend operations we can iteratively enlarge persegram till infinity and simultaneously create a sequence of equivalent persegrams (equivalence subclass). Since we want to find some simple characterization of equivalence, we should restrict our attention on the shortest possible representative of each such sequence to simplify the following lemmata and make the whole theory more lucid.

In other words, it means that we should consider persegrams on which no reduce operation can be applied only. Since *bullets extension/reduction* is a special part of the *subset extension/reduction* we may define so called *reduced persegram* as a persegram on which *subset reduction* can not be applied.

4 Conclusion

In this paper we gave a short introduction into equivalence problem. This problem includes several sub-problems where one of them is how to recognize whether two given persegrams are equivalent "on the first sight". The solution to this problem could be a direct characterization involving some invariants sufficient for equivalence.

We reminded two invariant properties introduced earlier: *Connections set* and *Ordering conditions*. The question whether these invariants are sufficient to decide of equivalence of given persegrams remains open.

We introduced four different operations (IE operations) on persegrams which preserve induced independence model. Combined together, they are able to generate the class of equivalent persegrams. That characterization is indirect in the following sense: If two persegrams can be transformed from one to the other by a sequence of IE operations, then the persegrams are independence equivalent. Anyway, this characterization is indirect in the sense that, if two persegrams over same set of variables are given, then searching of such a sequence can be time demanding or even impossible. However, indirect characterization offers a method to generate a class of equivalent persegrams. The following questions remain open: Can IE-operations generate the whole class of equivalent persegrams? Does there always exist a sequence of IE-operations transforming one persegram into another equivalent?

We also introduced some motivation that could be helpful for the following research. We should restrict our attention to some representative of every equivalence subclass. We hope, it makes answering the previous questions easier.

References

[1] R. Jiroušek. *Multidimensional Compositional Models.* Preprint DAR - ÚTIA 2006/4, ÚTIA AV ČR, Prague, (2006).

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

[3] M. Studený. *O strukturách podmíněné nezávislosti.* Rukopis série přednášek. Prague (2008).

[4] R. Merris: Graph Theory. Wiley Interscience, New York 2001.

[5] V. Kratochvíl. *Equivalence Problem in Compositional Models.* Doktorandské dny 2008, Nakladatelství ČVUT, Praha, p.125-134, 2008.

[6] V. Kratochvíl. *Equivalence Problem in Persegrams.* WUPES 2009, In preparation.