Conditioning in Compositional models

Václav Kratochvíl

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

Abstract. Reasoning by cases or assumptions is a common form of human reasoning. In case of probability reasoning, this is modeled by conditioning of a multidimensional probability distribution. Compositional models are defined as a multidimensional distributions assembled from a sequence of low-dimensional probability distributions, with the help of operators of composition. We call this sequence its generating sequence.
In case of compositional models, the conditioning process can be viewed as a transformation of one generating sequence into another one - preferably with the smallest number of local changes. It appears that the conditioning process is simple when conditioning variable appears in the argument of the first distribution of the corresponding generating sequence. That is why we introduce the so called flexible sequences. Flexible sequences are those, which can be reordered in many ways that each variable can appears among arguments of the first distribution. In this paper, we study the problem of flexibility in light of the very recent complex solution of the equivalence problem. Note that by the equivalence problem we understand how to recognize whether two generating sequence structures induce the same set of conditional independence assertions - the so called induced independence models.

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 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.

Reasoning by cases or assumptions is a common form of human reasoning. In case of probability reasoning, this is modeled by conditioning of a multidimensional probability distribution. Let us introduce the problem of conditioning a distribution that is represented in a form of a compositional model. Generally, conditioning process can be viewed as a transformation of one probability distribution into another. When representing a distribution in a form of a compositional model, we understand conditioning as a transformation of its generating sequence into another one - preferably with the smallest number of *local changes* (inspired by Lauritzen-Spiegelhalter's local computations). By a local change we understand either a change of just one distribution from the corresponding generating sequence (its recalculation), or permutation of the generating sequence.

The conditioning problem was briefly discussed in [1]. In the very same publication, there was also given an example illustrating conditioning in a simple distribution $\pi(u, v, w)$ represented by a compositional model with a generating sequence $\pi_1(u, v), \pi_2(v, w)$. There was also stated a theorem how to deal with the case when conditioning variable appears in the argument of the first distribution of the corresponding generating sequence - Assertion 1 in here - as well as concept of *flexible sequences*. We further investigate flexible sequences in this text - primarily using new evidence about independence equivalent permutations [6] of generating sequences.

1.1 Notation and Basic Properties

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,

⋆ 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.

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.

To compose low-dimensional distributions and to get a distribution of a higher dimension, we use the so-called operator of composition. It is described in the following definition:

Definition 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) $\forall x \in \times_{j \in K} \mathbf{X}_j; (\pi_2(x) = 0 \Longrightarrow \pi_1(x) = 0)$. 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 $\frac{0 \cdot 0}{0} = 0$*

1.2 Generating sequence

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, and we denote the distribution represented by a generating sequence $\pi_1, \pi_2, \ldots, \pi_n$ as

$$\triangleright_{\pi_1, \pi_2, \pi_3 \ldots, \pi_n} = (\ldots((\pi_1 \triangleright \pi_2) \triangleright \pi_3) \triangleright \ldots \triangleright \pi_{n-1}) \triangleright \pi_n.$$

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* .

1.3 Structure

Consider a compositional model defined by a generating sequence $\pi_1(U_1), \pi_2(U_2), \ldots, \pi_n(U_n)$. Then the sequence of sets $U_1, U_2, \ldots, U_n$ is called *model structure* and it is usually denoted by symbol $\mathcal{P}$. If not specified otherwise, $\mathcal{P} = U_1, \ldots, U_n$ where $(U_1 \cup \ldots \cup U_n) = N$, and we say that $\mathcal{P}$ is defined over N and $U_i \in \mathcal{P}$ for every $i \in \{1, \ldots, n\}$. Sets defining the structure are called *columns* to distinguish them from general sets of variables. Moreover we recognize the auxiliary sets $K_i^{\mathcal{P}}$ which reflects the ordering in $\mathcal{P}$ - $K_i^{\mathcal{P}}$ is the i-th set from $\mathcal{P}$; e.g. for $\mathcal{P} = U_1, \ldots, U_n$ holds that $K_i^{\mathcal{P}} = U_i$ for all $i = 1, \ldots, n$.

The reason for double notation of the same set from a structure is the following: Consider a situation when $\mathcal{P} = U_1, U_2, U_3, U_4$ and let $\mathcal{P}'$ be its permutation such that $\mathcal{P}' = U_3, U_1, U_4, U_2$ for example. Then U_3 is the first column in sequence $\mathcal{P}'$ and U_1 is the second column in $\mathcal{P}'$. This can be easily expressed as $U_3 \equiv K_1^{\mathcal{P}'}$ and $U_1 \equiv K_2^{\mathcal{P}'}$ now.

In addition, each column $K_i^{\mathcal{P}}$ can be divided into two disjoint parts with respect to the structure. We denote them $R_i^{\mathcal{P}}$ and $S_i^{\mathcal{P}}$, where $R_1^{\mathcal{P}} = K_1^{\mathcal{P}}$ and $R_i^{\mathcal{P}} = K_i^{\mathcal{P}} \setminus (K_1^{\mathcal{P}} \cup \ldots \cup K_{i-1}^{\mathcal{P}})$ $\forall i = \{2, \ldots, |\mathcal{P}|\}$. $S_1^{\mathcal{P}} = \emptyset$ and $S_i^{\mathcal{P}} = K_i^{\mathcal{P}} \cap (K_1^{\mathcal{P}} \cup \ldots \cup K_{i-1}^{\mathcal{P}})$ $\forall i = \{2, \ldots, |\mathcal{P}|\}$.

It has the following meaning: $R_i^{\mathcal{P}}$ denotes the variables first occurring in the i-th column of the sequence $\mathcal{P}$ (meaning from left to right). Conversely, $S_i^{\mathcal{P}}$ denotes variables from i-th set of $\mathcal{P}$ which have already been used in some foregoing column. Observe that $K_i^{\mathcal{P}} = R_i^{\mathcal{P}} \cup S_i^{\mathcal{P}}$. Columns not inducing any new variables into the sequence are called *trivial* (i.e. $K_i^{\mathcal{P}} \in \mathcal{P} : R_i^{\mathcal{P}} = \emptyset$). The super index $\mathcal{P}$ may be omitted if the context is clear. $|\mathcal{P}|$ denotes the number of sets in the structure, i.e., $|\mathcal{P}| = n$ for $\mathcal{P} = U_1, \ldots, U_n$.

Example 1. For a generating sequence $\pi_1(u)$, $\pi_2(v, w)$, $\pi_3(u, v, x)$, $\pi_4(w, x, y)$, $\pi_5(x, y, z)$, its structure is $\mathcal{P} = \{u\}, \{v, w\}, \{u, v, x\}, \{w, x, y\}, \{w, y, z\}$ and $|\mathcal{P}| = 5$.

$K_1^{\mathcal{P}} = u$ $\quad R_1^{\mathcal{P}} = u$ $\quad S_1^{\mathcal{P}} = \emptyset$
$K_2^{\mathcal{P}} = \{v, w\}$ $\quad R_2^{\mathcal{P}} = \{v, w\}$ $\quad S_2^{\mathcal{P}} = \emptyset$
$K_3^{\mathcal{P}} = \{u, v, x\}$ $\quad R_3^{\mathcal{P}} = x$ $\quad S_3^{\mathcal{P}} = \{u, v\}$
$K_4^{\mathcal{P}} = \{w, x, y\}$ $\quad R_4^{\mathcal{P}} = y$ $\quad S_4^{\mathcal{P}} = \{w, x\}$
$K_5^{\mathcal{P}} = \{w, y, z\}$ $\quad R_5^{\mathcal{P}} = z$ $\quad S_5^{\mathcal{P}} = \{x, y\}$

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

$$]\cdot[_{\mathcal{P}} \colon 2^N \to \{1, \ldots, |\mathcal{P}|\}$$

such that for fixed structure $\mathcal{P}$, $U \subseteq N$, $]U[_{\mathcal{P}} = max_{u \in U}\{i : u \in R_i^{\mathcal{P}}\}$. Hence $]U[_{\mathcal{P}}$ equals the maximal index i such that $u \in U$ and $u \in R_i^{\mathcal{P}}$. Due to the previously established notation, it can be said that $K_{]u[_{\mathcal{P}}}^{\mathcal{P}}$ is a column $K_i^{\mathcal{P}}$ for which $u \in R_i^{\mathcal{P}}$, i.e., $]u[_{\mathcal{P}} = i : u \in R_i^{\mathcal{P}}$. The symbol $\mathcal{P}$ may be omitted in $]u[_{\mathcal{P}}$ if the context is clear - for example when dealing with one structure only.

Example 2. Consider structure $\mathcal{P}$ from Example 1. One can read the following properties: $]u[= 1$, $]\{u, v\}[= 2$, $]\{u, w\}[= 2$, $]x[= 3$, $]y[= 4$, $]\{u, v, w, x, y, z\}[= 5$, and $]z[= 5$. Similarly to Example 1

$K_{]u[}^{\mathcal{P}} = u$ $\quad R_{]u[}^{\mathcal{P}} = u$ $\quad S_{]u[}^{\mathcal{P}} = \emptyset$
$K_{]v[}^{\mathcal{P}} = \{v, w\}$ $\quad R_{]w[}^{\mathcal{P}} = \{v, w\}$ $\quad S_{]\{v,w\}[}^{\mathcal{P}} = \emptyset$
etc.

Definition 2. *For a structure $\mathcal{P}$ over N we introduce a binary relation $\preceq_{\mathcal{P}}$ such that for two non-empty sets of variables $U, V \subseteq N$: $U \preceq_{\mathcal{P}} V$ if and only if $]U[_{\mathcal{P}} \leq]V[_{\mathcal{P}}$. Moreover, we introduce its strict version $\prec_{\mathcal{P}}$: $U \prec_{\mathcal{P}} V$ if and only if $]U[_{\mathcal{P}} <]V[_{\mathcal{P}}$.*

The symbol $\mathcal{P}$ may be omitted in $\prec_{\mathcal{P}}$ and $\preceq_{\mathcal{P}}$ if the context is clear.

Example 3. Consider the structure from Example 1 again. According to the former definition one can see that $u \prec v \preceq w \prec x \prec y \prec z$ in that structure. Similarly, for subsets $\{u, v, w\} \prec z$, $\{u, v\} \preceq w$, $\{u, x\} \prec y$, etc.

1.4 Permutations

To increase the clarity of the text, the standard concept of permutation (including the notation) is used. In mathematics, the notion of permutation is used with several slightly different meanings, all related to the act of permuting (rearranging in an ordered fashion) objects or values. Informally, a permutation of a set of values is an arrangement of those values into a particular order. Thus there are six permutations of the set $\{1, 2, 3\}$, namely $[1, 2, 3]$, $[1, 3, 2]$, $[2, 1, 3]$, $[2, 3, 1]$, $[3, 1, 2]$, and $[3, 2, 1]$.

Definition 3. *Let X be a set, then a* permutation *of X is a bijection $\sigma : X \to X$.*

There are several ways how to write permutations. We use the method of *product of disjoint cycles*, since it is well known that one can express every permutation as a product of disjoint cycles. Let $i_1, i_2, \ldots, i_r$ be distinct elements of X. The *r-cycle* $(i_1\ i_2\ \ldots\ i_r)$ is the permutation which maps $i_1 \mapsto i_2$, $i_2 \mapsto i_3$, $\ldots$, $i_{r-1} \mapsto i_r$, $i_r \mapsto i_1$ and fixes all other points in X. For a permutation of n symbols, the collection of all permutation of this set is denoted by T_n.

If σ is a permutation of the set X, we shall write $i\sigma$ for the image of the element $i \in X$ under σ (rather then $\sigma(i)$). The principal reason for doing this is that it makes composition of permutations much easier: $\sigma_1\sigma_2$ will mean apply σ_1 first and then apply σ_2 rather than the other way around.

In case of generating sequences, or sequences of probability distributions generally, we permute elements ordering. I.e. In a cycle such as $\sigma = (1\ 2\ 4\ 5)$ we mean that the permutation maps $1st$ to $2nd$, maps $2nd$ to $4th$, maps $4th$ to $5th$ and $5th$ to $1st$ one. Hence for $\pi_1, \ldots, \pi_5$, the sequence $(\pi_1, \ldots, \pi_5)\sigma \equiv \pi_1, \pi_5, \pi_3, \pi_2, \pi_4$. Similarly for structures: Let $\mathcal{P} = U_1, U_2, \ldots, U_8$ be a compositional model structure and $\sigma \in T_8$ be a permutation. In case of cycle $\sigma = (1\ 2\ 4\ 5)$ holds that $\mathcal{P}\sigma = U_5, U_1, U_3, U_2, U_4, U_6, U_7, U_8$.

The product of two permutations σ_1 and σ_2 is the function obtained by applying σ_1 first and then applying σ_2. Since we are writing maps on the right, we denote this by $\sigma_1\sigma_2$. Note that in general $\sigma_1\sigma_2 \neq \sigma_2\sigma_1$.

We give a special accent for the following very short cycles - *transpositions*: A *transposition* is a cycle of length two (that is, with two elements) - the so called 2-cycle. Thus a transposition is a permutation

$(i\ j)$ which simply swaps round the two elements i and j. Transpositions are useful for the following reason: Is is well known that *every permutation can be expressed as a product of transpositions.*

For example,

$$(1\ 2\ 3\ 4\ 5) = (1\ 2)(1\ 3)(1\ 4)(1\ 5)$$

does what we want for a cycle of length 5. Analogous calculations establish the same for other lengths. A *permutation* of a given compositional model is a permutation of its generating sequence.

2 Conditioning

In [1] were illustrated a general fact: *computation of a conditional distribution $\pi(\cdot|u = x_u)$ (for $\pi = \pi_1 \triangleright \ldots \triangleright \pi_n$) is easy only if the conditioning variable u appears among arguments of the first distribution π_1.* In the following theorem, that was originally published in [1], one can see that the process of conditioning is simple in case that the conditioning variable appears among arguments of the first distribution in the respective generating sequence.

Assertion 1 *Let $\pi_1, \pi_2, \ldots, \pi_n$ be a generating sequence with structure $\mathcal{P}$ over N and $u \in K_1^{\mathcal{P}}$. Then, for any value x_u of the variable u, for which $\pi_1(u = x_u) > 0$,*

$$(\pi_1 \triangleright \pi_2 \triangleright \ldots \triangleright \pi_n)\left((K_1^{\mathcal{P}} \cup \ldots \cup K_n^{\mathcal{P}}) \setminus u | u = x_u\right) = \kappa_1 \triangleright \kappa_2 \triangleright \ldots \triangleright \kappa_n,$$

where for all $i = 1, 2, \ldots, n$

$$\kappa_i(K_i^{\mathcal{P}} \setminus u) = \begin{cases} \pi_i(K_i^{\mathcal{P}}) & \text{if } u \notin K_i^{\mathcal{P}} \\ \pi_i(K_i^{\mathcal{P}} | u = x_u) & \text{if } u \in K_i^{\mathcal{P}}. \end{cases}$$

In light of Assertion 1, it seems reasonable to study the raised question: in which case we may reorder the generating sequence in a way that a desired variable appears among arguments of its first probability distribution. To specify the problem, we define the so called *flexible sequences.* Note, that the concept of a generating sequence *flexibility* was originally introduced in [1] during studies of a stronger property - the so called *decomposability.*

Definition 4. *A generating sequence $\pi_1, \pi_2, \ldots, \pi_n$ with structure $\mathcal{P}$ is called* flexible *if for all $u \in K_1^{\mathcal{P}} \cup \ldots \cup K_n^{\mathcal{P}}$ there exists a permutation $\sigma \in T_n$ such that $u \in K_1^{\mathcal{P}\sigma}$ and*

$$\triangleright_{(\pi_1, \pi_2, \ldots, \pi_n)\sigma} = \triangleright_{\pi_1, \pi_2, \ldots, \pi_n}.$$

In other words, flexible sequences are those, which can be reordered in many ways so that each variable can appears among arguments of the first distribution. However, it does not mean, that each distribution appears at the beginning of the generating sequence. (If this would be the case, then flexible sequence would be a subclass of special - the so called *perfect* [1] - sequences.

Observe that the problem of conditioning by a variable turns into a problem of *flexibility* in light of Assertion 1. We resign to study of other possible conditioning algorithms and we demand the conditional variable among the arguments of the first distribution of the generating sequence.

Every structure induce a system of conditional independence assertions valid for every multidimensional probability distribution represented by a compositional model. We call this system *induced independence model* [3, 5] and for structure $\mathcal{P}$ it is denoted by $\mathcal{I}(\mathcal{P})$. It seems reasonable to focus on those permutations only, that do not affect this system of independence assertions - i.e. structures of respective sequences are independence equivalent. Imagine that we omit this requirement. Then, during a permutation, structure loose its power to guarantee validity of independence e.g. $u \perp\!\!\!\perp v | Z$. Then, to represent the same probability distribution by both - original as well as permutated sequence - the independence has to be induced by numerical properties of distributions from the generating sequence and vice versa. To check all such conditions would be extremely time-demanding. Hence we will study only those sequences, where their flexibility is guaranteed mainly by their structures.

3 Flexible structure

To explore our approximation to flexible sequences from above (where we restrict ourselves to those flexible sequences that induce simultaneously independence equivalent structures only), we will define the so called *flexible structure.*

Definition 5. *Let $\mathcal{P}$ be a structure over N. $\mathcal{P}$ is* flexible *if $\forall u \in N$ exists a permutation σ such that $u \in K_1^{\mathcal{P}\sigma}$ and $\mathcal{I}(\mathcal{P}) = \mathcal{I}(\mathcal{P}\sigma)$.*

The problem of independence equivalent structures was well examined in [3, 5, 6]. Recall, that there exists several characterizations of independence equivalence. In this context let us highlight the so called *F-condition set* $\mathcal{F}(\mathcal{P})$ - one of independence equivalence invariants. First, we will show auxiliary property of *connection.* See [5] for more about both these properties.

Definition 6. *Consider structure $\mathcal{P}$ over N and two distinct variables $u, v \in N$. We say that u, v are* connected *in $\mathcal{P}$ - $u \leftrightarrow_{\mathcal{P}} v$ if $u \in K^{\mathcal{P}}_{]v[} \vee v \in K^{\mathcal{P}}_{]u[}$. Otherwise they are not connected $u \not\leftrightarrow_{\mathcal{P}} v$.*

Definition 7. *Consider a structure $\mathcal{P}$ over N and three disjoint variables $u, v, w \in N$. We say that the triplet $\langle u, v|w\rangle$ forms* F-condition *if*

(a) $\{u, v\} \prec_{\mathcal{P}} w$,
(b) $u \leftrightarrow_{\mathcal{P}} w \wedge v \leftrightarrow_{\mathcal{P}} w$,
(c) $u \not\leftrightarrow_{\mathcal{P}} v$.

The set of triples $\mathcal{F}(\mathcal{P}) = \{\langle u, v|w\rangle : \{u, v\} \prec_{\mathcal{P}} w \wedge \{u, v\} \leftrightarrow_{\mathcal{P}} w \wedge u \not\leftrightarrow_{\mathcal{P}} v\}$ is called F-condition set *induced by $\mathcal{P}$.*

Assertion 2 *For two structure $\mathcal{P}, \mathcal{P}'$ over the same set of variables N holds*

$$\mathcal{I}(\mathcal{P}) = \mathcal{I}(\mathcal{P}') \Rightarrow \mathcal{F}(\mathcal{P}) = \mathcal{F}(\mathcal{P}').$$

Non-emptiness of the induced F-condition set $\mathcal{F}(\mathcal{P})$ has a major impact on the flexibility of the respective structure. Judge for yourself: If the disjoint triplet $\langle u, v|w\rangle \in \mathcal{F}(\mathcal{P})$ then by (a) from Definition 7 and the fact that $\mathcal{F}(\mathcal{P})$ is one of independence class invariants - Assertion 2, it follows that there is no structure $\mathcal{P}'$ equivalent with $\mathcal{P}$ such that $w \in K_1^{\mathcal{P}'}$. Indeed, there always have to be some foregoing columns introducing u and v first in $\mathcal{P}'$. Moreover, (c) implies that if $\langle u, v|w\rangle \in \mathcal{F}(\mathcal{P})$ then variable w may appear not earlier than in the third column of any independence equivalent structure for the first time (both u and v are introduced in different sets). Hence the emptiness of $\mathcal{F}(\mathcal{P})$ is a necessary condition for flexibility of $\mathcal{P}$:

Remark 1. If a structure $\mathcal{P}$ is flexible, then $\mathcal{F}(\mathcal{P}) = \emptyset$.

3.1 Column covering

In this section we shall see that the emptiness of $\mathcal{F}(\mathcal{P})$ is not only necessary but also sufficient condition for the respective structure $\mathcal{P}$ flexibility. To prove this we need to employ the specific consequence of $\mathcal{F}(\mathcal{P}) = \emptyset$. This consequence deals with specific structure behavior regarding its any non-trivial column - the so called *column covering.*

Definition 8. *Consider structure $\mathcal{P}$ over N and variable $u \in N$. Column $K^{\mathcal{P}}_{]u[}$ is* covered *in structure $\mathcal{P}$ if either $]u[_{\mathcal{P}}= 1$ or if there exists variable $v \in N$: $v \prec_{\mathcal{P}} u$ such that $S^{\mathcal{P}}_{]u[} \subseteq K^{\mathcal{P}}_{]v[}$. We say that $K^{\mathcal{P}}_{]v[}$ is* covering column *of $K^{\mathcal{P}}_{]u[}$.*

The essential feature of $\mathcal{F}(\mathcal{P}) = \emptyset$ lies in the fact formulated in the following lemma:

Lemma 1. *Let $\mathcal{P}$ be a structure such that $\mathcal{F}(\mathcal{P}) = \emptyset$. Then every its non-trivial column is covered in it.*

Proof. Note, that the first and second column of any structure $\mathcal{P}$ is covered by definition: Indeed, while $K_1^{\mathcal{P}}$ is covered by Definition 8; for $K_2^{\mathcal{P}}$ holds that $S_2^{\mathcal{P}} = K_1^{\mathcal{P}} \cap K_2^{\mathcal{P}} \subseteq K_1^{\mathcal{P}}$ and $K_1^{\mathcal{P}}$ is a covering column of $K_2^{\mathcal{P}}$. (Note that $K_1^{\mathcal{P}}$ is non-trivial by its definition in any structure $\mathcal{P}$.)

Choose an arbitrary $w \in N$ such that $]w[_{\mathcal{P}}\geq 3$. One can distinguish two cases:

I. $|S^{\mathcal{P}}_{]w[}| \leq 1$
II. $|S^{\mathcal{P}}_{]w[}| \geq 2$

In case of $S^{\mathcal{P}}_{]w[} \leq 1$, either $S^{\mathcal{P}}_{]w[} = \emptyset$, and then $K_1^{\mathcal{P}}$ can be its covering column, or $|S^{\mathcal{P}}_{]w[}| = 1$. Put $u = S^{\mathcal{P}}_{]w[}$ and observe that $K^{\mathcal{P}}_{]u[}$ is its covering column by Definition 8.

Assume that $|S^{\mathcal{P}}_{]w[}| \geq 2$. Choose and fix $v \in S^{\mathcal{P}}_{]w[}$ such that $v \succeq_{\mathcal{P}} v'$ for all other $v' \in S^{\mathcal{P}}_{]w[}$. Now, let us show that $S^{\mathcal{P}}_{]w[} \subseteq K^{\mathcal{P}}_{]v[}$ by considering the opposite for a contradiction - i.e. let $\exists u \in S^{\mathcal{P}}_{]w[}$ such that $u \notin K^{\mathcal{P}}_{]v[}$. Then $u \prec_{\mathcal{P}} v$ by the choice of v and $u \not\leftrightarrow_{\mathcal{P}} v$ by Definition 6. It implies that $\langle u, v|w\rangle \in \mathcal{F}(\mathcal{P})$ by Definition 7 which contradicts with the lemma supposition $\mathcal{F}(\mathcal{P}) = \emptyset$. Hence $S^{\mathcal{P}}_{]w[} \subseteq K^{\mathcal{P}}_{]v[}$ and $K^{\mathcal{P}}_{]w[}$ is covered by $K^{\mathcal{P}}_{]v[}$ in this case, which finishes the proof.

3.2 Independence equivalent permutations

The notion of covering column has a close connection to the so called *constant transposition* [4] and related *left cycle permutation* [6]. For the purpose of this text, we define them in a little but different way than in [4, 6].

Definition 9. *Consider structure $\mathcal{P}$, $i \in \{1, \ldots, |\mathcal{P}| - 1\}$, $k \in \{2, \ldots, |\mathcal{P}| - i\}$ such that $K_i^{\mathcal{P}} \supseteq S_{i+k}^{\mathcal{P}}$. Then we call a cycle $\sigma_L = (i{+}1\ i{+}2\ \ldots\ i{+}k)$ a* left cycle permutation *in $\mathcal{P}$. We say that $\mathcal{P}\sigma_L$ is left cycle permutation of $\mathcal{P}$.*

Assertion 3 *If $\mathcal{P}\sigma_L$ is* left cycle permutation *of $\mathcal{P}$ then $\mathcal{I}(\mathcal{P}) = \mathcal{I}(\mathcal{P}\sigma_L)$.*

In the following we will employ also the following simple assertion from [4]:

Assertion 4 *For every structure $\mathcal{P}$ such that $|\mathcal{P}| \geq 2$ and $\sigma = (1\ 2)$ holds that $\mathcal{I}(\mathcal{P}) = \mathcal{I}(\mathcal{P}\sigma)$.*

More precisely, by definition of *left cycle permutation* - Definition 9, any covered column may be moved just behind its covering column using left cycle permutation. Indeed, let $K_i^{\mathcal{P}}, K_{i+k}^{\mathcal{P}}$ be a couple of covering and covered column - i.e. $K_i^{\mathcal{P}} \supseteq S_{i+k}^{\mathcal{P}}$ by Definition 8. Then $(i{+}1\quad i{+}2\quad \ldots\quad i{+}k)$ is a left cycle permutation in $\mathcal{P}$. Using this and the fact that transposition $(1\ 2)$ preserves induced independence model in any structure $\mathcal{P}$, one can easily proof that in case of $\mathcal{F}(\mathcal{P}) = \emptyset$, any non-trivial column may be moved to the first position in the structure while the induced independence model is the same.

Lemma 2. *For structure $\mathcal{P}$ with $\mathcal{F}(\mathcal{P}) = \emptyset$ and every column $K_{]u[}^{\mathcal{P}}$, there exists a permutation $\sigma \in T_{|\mathcal{P}|}$ such that $(]u[_{\mathcal{P}})\sigma = 1$ ($K_{]u[}^{\mathcal{P}} \equiv K_1^{\mathcal{P}\sigma}$) and σ may be internally replaced by a sequence of left cycle permutations and $(1\ 2)$ transpositions.*

Proof. Let us proof the assertion of this lemma by induction on $i \in \{1, \ldots, |\mathcal{P}|\}$. The induction hypothesis for $i \geq 1$ is that there exists a permutation σ such that $i\sigma = 1$, $j\sigma = j$ for all $i < j \leq |\mathcal{P}|$, and σ may be replaced by a sequence of left cycle permutations and $(1\ 2)$ transpositions. It is evident for $i = 1$. σ is *identical permutation* in this case.

Assume $i =]u[_{\mathcal{P}} \geq 2$ and that the implication holds every $v \in N$ such that $]v[_{\mathcal{P}} <]u[_{\mathcal{P}}$. Then $K_i^{\mathcal{P}} \equiv K_{]u[}^{\mathcal{P}}$ is covered by Lemma 1 - i.e. $\exists K_{]v[}^{\mathcal{P}}$ such that

$$S_{]u[}^{\mathcal{P}} \subseteq K_{]v[}^{\mathcal{P}}. \tag{1}$$

Since $]v[_{\mathcal{P}} <]u[_{\mathcal{P}}$ then by induction hypothesis there exists $\sigma \in T_{|\mathcal{P}|}$ such that

$$K_1^{\mathcal{P}\sigma} = K_{]v[}^{\mathcal{P}} \tag{2}$$

and $j\sigma = j$ for all $]v[_{\mathcal{P}} < j \leq |\mathcal{P}|$. Observe that then however

$$S_j^{\mathcal{P}} = S_j^{\mathcal{P}\sigma} \text{ for all } j >]v[_{\mathcal{P}} \tag{3}$$

by definition of $S_{]v[}^{\cdot}$. Combining all the expressions (1), (2), and (3) with the fact that $]u[_{\mathcal{P}} >]v[_{\mathcal{P}}$, we can easily obtain the relationship $S_{]u[}^{\mathcal{P}\sigma} \subseteq K_1^{\mathcal{P}\sigma}$ guaranteing that $\sigma_L = (2\ 3\ \ldots\ i)$ is left cycle permutation in $\mathcal{P}\sigma$. Put $\sigma_{cb} = (1\ 2)$. Then $\sigma' = \sigma\sigma_L\sigma_{cb}$ guarantees that $i\sigma' = 1$, $j\sigma = j$ for all $i < j \leq |\mathcal{P}|$ by their definition. Note that σ may be replaced by sequence of left cycle permutations and $(1\ 2)$ transpositions by induction hypothesis. Then the lemma is proved.

Employing Assertions 5 and 6, one can easily conclude:

Corollary 1. *For structure $\mathcal{P}$ with $\mathcal{F}(\mathcal{P}) = \emptyset$ and every column $K_{]u[}^{\mathcal{P}}$, there exists a permutation $\sigma \in T_{|\mathcal{P}|}$ such that $(]u[_{\mathcal{P}})\sigma = 1$ ($K_{]u[}^{\mathcal{P}} \equiv K_1^{\mathcal{P}\sigma}$) and $\mathcal{I}(\mathcal{P}) = \mathcal{I}(\mathcal{P}\sigma)$.*

Realizing the fact that for structure $\mathcal{P}$ over N and every variable $u \in N$ there exists column $K_{]u[}^{\mathcal{P}}$, it follows that $\mathcal{F}(\mathcal{P}) = \emptyset$ is sufficient condition for flexibility of the respective structure $\mathcal{P}$. Hence, using Remark 1, emptiness of $\mathcal{F}(\mathcal{P})$ is not only necessary but also sufficient condition for flexibility of structure $\mathcal{P}$:

Corollary 2. *Structure $\mathcal{P}$ is flexible $\Leftrightarrow \mathcal{F}(\mathcal{P}) = \emptyset$.*

Notice that condition of $\mathcal{F}(\mathcal{P}) = \emptyset$ means the existence of equivalent permutation of $\mathcal{P}$ for every variable where the variable appears in its first column of the permutation. Hence, there could be non-trivial columns from $\mathcal{P}$ that do not appear at the beginning of any such an equivalent structure. On the contrary, $\mathcal{F}(\mathcal{P}) = \emptyset$ guarantees the existence of equivalent permutation for every non-trivial column such this column appears at the beginning of the equivalent permutation.

In another words, it allows us to move any non-trivial column to the sequence beginning using some independence equivalent permutation. See the following example to illustrate the difference:

Example 4. Consider $U_1 = \{u, v\}, U_2 = \{v, w\}$, and $U_3 = \{w, x\}$. Observe that for structures U_1, U_2, U_3 and U_3, U_2, U_1, the corresponding formal ratios coincide:

$$U_1, U_2, U_3 : \ \frac{\{u, v\} \cdot \{v, w\} \cdot \{w, x\}}{v \cdot w},$$

$$U_3, U_2, U_1 : \ \frac{\{u, v\} \cdot \{v, w\} \cdot \{w, x\}}{v \cdot w}.$$

Then structures U_1, U_2, U_3 and U_3, U_2, U_1 are independence equivalent - formal ratio is one of characterizations of independence equivalence [6]. Moreover, since any of u, v, w, and x appears in the first column of one of those structures, then this is sufficient for flexibility of U_1, U_2, U_3.

Note that since $\mathcal{F}(U_1, U_2, U_3) = \emptyset$, then there exists also an equivalent permutation with U_2 in the first position by Lemma 2. (specifically U_2, U_1, U_3). However, its existence is not necessary for flexibility of U_1, U_2, U_3 in this case.

3.3 Flexible structures versus flexible sequences

For a generating sequence and its independence equivalent permutation (corresponding structures are independence equivalent), the pairwise consistency of considered distributions guarantees that both sequences are equivalent simultaneously (they both represents identical multidimensional distribution) [6]. Considering definition of a structure flexibility, flexible structures are closely connected with independence equivalence and thus with IE operations [4] as well. The impact of IE operations on arbitrary distribution represented by a compositional model was described in [6]. We will recall two needed assertions:

Assertion 5 *Let $\pi_1, \pi_2, \ldots, \pi_n$ be a generating sequence with structure $\mathcal{P}$. If σ_L is* left cycle permutation *in $\mathcal{P}$ then*

$$\triangleright_{\pi_1, \pi_2, \ldots, \pi_n} = \triangleright_{(\pi_1, \pi_2, \ldots, \pi_n)\sigma_L}.$$

Assertion 6 *Let $\pi_1, \pi_2, \ldots, \pi_n$ be a generating sequence and $\sigma = (1\ 2)$. If π_1, π_2 are consistent then*

$$\triangleright_{\pi_1, \pi_2, \ldots, \pi_n} = \triangleright_{(\pi_1, \pi_2, \ldots, \pi_n)\sigma}.$$

Considering the proof of Lemma 2, one can see that left cycle permutations and (1 2) transpositions were used only. Then, in case of pairwise consistency, one can end up with the following deduction:

Lemma 3. *If $\pi_1(U_1), \pi_2(U_2), \ldots, \pi_n(U_n)$ is a sequence of pairwise consistent probability distributions with flexible structure $U_1, U_2, \ldots, U_n$, then this sequence is flexible.*

Proof. This is a simple consequence of Corollary 2, Lemma 2, and iterative application of Assertions 5 and 6.

Remark 2. One may object, that the previous lemma - Lemma 3 - represents literally "a reinvention of a wheel". Indeed, note that the condition $\mathcal{F}(\mathcal{P}) = \emptyset$ corresponds to the so called *running intersection property (RIP)*. Its definition is the following: Let $U_1, U_2, \ldots, U_n$ is a sequence of sets. Then this sequence meets RIP if

$$\forall i = 2, \ldots, n \quad \exists j : (1 \leq j < i) \left(U_i \cap (\bigcup_{k=0}^{i-1} U_k) \subseteq U_j \right).$$

This definition can be easily rewritten in case that the sequence of sets represents a structure $\mathcal{P}$: Structure $\mathcal{P}$ meets RIP if

$$\forall i = 2, \ldots, |\mathcal{P}| \quad \exists j : (1 \leq j < i) \left(S_i^{\mathcal{P}} \subseteq K_j^{\mathcal{P}} \right).$$

I.e. in case of RIP, there is a covering column for each column (including trivial ones). In case when we do not consider trivial columns, then the condition of $\mathcal{F}(\mathcal{P}) = \emptyset$ coincides with RIP and one can find the following lemma in [1]: If $\pi_1(U_1), \pi_2(U_2), \ldots, \pi_n(U_n)$ is a sequence of pairwise consistent probability distributions such that $U_1, U_2, \ldots, U_n$ meets RIP then this sequence is flexible.

Considering the proof of Lemma 2, Lemma 3 may be slightly modified using relations to generating sequences given by Assertions 5 and 6. Recall, that while (1 2) transposition requires a consistency of respective distributions to guarantee the equality of the permutated compositional model, left cycle permutation has no additional claims on the corresponding distributions. Then, given the proof of Lemma 2, we do not need to require the pairwise consistency, we only need the consistency of those pairs of distributions that corresponds to covering-covered pairs of columns in the respective structure:

Corollary 3. *If $\pi_1(U_1), \pi_2(U_2), \ldots, \pi_n(U_n)$ is a sequence of probability distributions with flexible structure $U_1, U_2, \ldots, U_n$ such that those pairs of distributions corresponding to covering-covered pairs of columns from the structure are consistent, then this sequence is flexible.*

Note that this become handy in case of automatical checking of flexibility. *In case of a generating sequence of length n, we restrict the number of consistencies to verify from*

$$\frac{n(n-1)}{2}$$

to

$$n-1.$$

Note that consistency of "covered $\times$ covering" distributions ensures, in the case of structure flexibility, pairwise consistency. It is then (in the case of structure consistency) a sort of minimum spanning for pairwise consistency.

In case of structure, flexibility guarantees that for every non-trivial column exists an independence equivalent permutation with this column on the first position. Then, by generalization to generating sequence (with structure without non-trivial columns) it holds that for distribution from the sequence exists some equivalent permutation with the distribution at the beginning. This, although it has not been mentioned (see [2] for example), guarantees that every such a low-dimensional distribution represents marginal of the respective compositional model. Hence, those structures are *perfect.*

Conclusions

We give a brief introduction into the problem of flexible sequences in this short paper. In our approach we employ the recent solution of equivalence problem in the area of compositional models structures. The concept of flexible sequences is very strong and in case that we have a simple characterization of flexible sequences, the problem of conditioning would be simple. Briefly, flexible sequences are those, that may be reordered in many ways such that each variable appears among arguments of first distributions, and all permutated generating sequences represent the same multidimensional probability distribution. In this paper, we restrict ourselves to those flexible sequences whose structure induces the same system of conditional independence assertions as structures of all respective permutations. I.e. structures of all permutated sequences are independence equivalent. We have shown that in this case the necessary condition of generating sequence flexibility is emptiness of induced F-conditions set and consistency of several pairs of low-dimensional distributions from the sequence.

References

1. R. Jiroušek: *Multidimensional Compositional Models.* Preprint DAR - ÚTIA 2006/4, ÚTIA AV ČR, Prague, (2006).
2. R. Jiroušek: *Foundations of compositional model theory*, International Journal of General Systems - Volume 40, Issue 6, (2011), pp. 623-678.
3. V. Kratochvíl: *Equivalence Problem in Compositional Models*, WUPES'09, Eds: Kroupa T., Vejnarová J., WUPES'09, Liblice (2009).
4. V. Kratochvíl: *Motivatio for different characterization of Equivalent Persegrams*, Proceedings of the 12th Czech-Japan Seminar on Data Analysis and Decision Making under Uncertainty, Eds: Novák V., Pavliska V., Štěpnička M., Czech-Japan Seminar on Data Analysis and Decision-making under Uncertainty /12./, Litomyšl (2009).
5. V. Kratochvíl: *Characteristic Properties of Equivalent Structures in Compositional Models*, International Journal of Approximate Reasoning vol. 52,5 (2011), pp. 599-612.
6. V. Kratochvíl: *On Equivalence Problem in Compositional Models*, thesis for Ph.D. degree, in printing process. (2011)