

# Kybernetika

---

Vladislav Bína; Radim Jiroušek  
On computations with causal compositional models

*Kybernetika*, Vol. 51 (2015), No. 3, 525–539

Persistent URL: <http://dml.cz/dmlcz/144384>

## Terms of use:

© Institute of Information Theory and Automation AS CR, 2015

Institute of Mathematics of the Czech Academy of Sciences provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these *Terms of use*.



This document has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ: The Czech Digital Mathematics Library* <http://dml.cz>

## ON COMPUTATIONS WITH CAUSAL COMPOSITIONAL MODELS

VLADISLAV BÍNA AND RADIM JIROUŠEK

The knowledge of causal relations provides a possibility to perform predictions and helps to decide about the most reasonable actions aiming at the desired objectives. Although the causal reasoning appears to be natural for the human thinking, most of the traditional statistical methods fail to address this issue. One of the well-known methodologies correctly representing the relations of cause and effect is Pearl's causality approach.

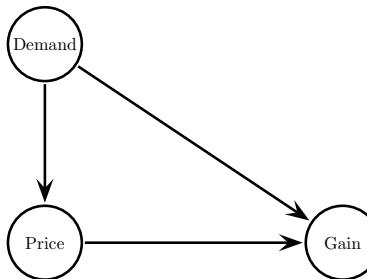
The paper brings an alternative, purely algebraic methodology of causal compositional models. It presents the properties of operator of composition, on which a general methodology is based that makes it possible to evaluate the causal effects of some external action. The proposed methodology is applied to four illustrative examples. They illustrate that the effect of intervention can in some cases be evaluated even when the model contains latent (unobservable) variables.

*Keywords:* causal model, conditioning, intervention, extension

*Classification:* 65C50, 97K50

### 1. INTRODUCTION

In many situations one would like to know what would happen if the current state of world is changed by an external intervention. For example, the governor of the national bank plans to intervene in favor (or, against the strength) of the national currency, and it would be nice to reveal all the consequences before such action is realized. Similar situations occur in everyday work of physicians: It would be nice to know in advance the effects of a specific treatment applied to a patient. To simplify our presentation, in this paper we will illustrate our considerations on a trivial example. Assume we trade in cement and have reliable suppliers, which enables us to control the price just on the basis of demand. Whenever the demand for cement increases we can slightly increase the price, and vice versa, which, as we believe, maximizes our profit. The situation is depicted in Figure 1. Imagine that a current demand for cement is on its minimum, we keep the price very low and therefore our profit is almost negligible. The question is whether we can afford to increase the price of cement regardless of the low current demand. What will be the impact of such an action to our profit.



**Fig. 1.** An example of causal relations (cement trading).

It is well known that the questions asked in the previous paragraph cannot be answered solely with the help of statistical data [2]. Such data may serve to answer the question: What was our profit in case that the price of cement was high? But now, we are asking the question: What will be our profit if we set the price of cement high? Naturally, in the past when the price was high, the demand was high, too, and therefore also the profit was high. Now, we cannot expect the increase of the price will increase the demand, rather the opposite. Therefore, such a direct utilization of the statistical data collected in the past is useless.

Nevertheless, when studying the book [7] by Judea Pearl we can see that such data can help us for causal models construction. Naturally, such data will not help us to find out causal relations. They can help us to reveal a dependence of variables but not the direction of causality. This must be done on the basis of some other knowledge, and the data can be used just to estimate the necessary parameters, the necessary probabilities. For a more detailed introduction to the causal inference and interesting examples of application in (strategic) management see, e. g., book by Michael Ryall and Aaron Bramson [8].

Most of the tools for causal model description use directed graphs to express the (asymmetric) relations of cause and effect (for an example see [1]). It also holds for Pearl's causal networks [7]. In contrast to this common approach, in this paper we take advantage of the algebraic apparatus of compositional models [3]. In fact, this paper is a continuation of the contribution to IPMU 2014 conference [4], where the first ideas on causal compositional models were published. The present paper is organized as follows. Section 2 recalls the main operator, which is used to assemble compositional models, so called operator of composition, and its properties. In Section 3, we show difference between conditioning and intervention, define causal compositional models and present general formulae for the computation of intervention in causal compositional models. The examples illustrating both the application of these formulae and the situations, in which they are of no help can be found in Section 4.

## 2. BASIC NOTIONS AND NOTATION

In this paper we consider a set of  $n$  finite valued *variables*  $\{X_1, X_2, \dots, X_n\}$ , subsets of which are denoted by lower-case Roman alphabets (e. g.,  $x$ ,  $y$ , and  $z$ ).  $\langle X_i \rangle$  denotes

the set of values (states) of variable  $X_i$ , and analogously, for sets of variables  $x, y$  the respective Cartesian products of all combinations of their values are denoted by  $\langle x \rangle, \langle y \rangle$ , respectively. Elements of these sets, i.e., the (combinations of) values of variables will be denoted by lower-case boldface Roman characters (e.g.,  $\mathbf{a} \in \langle x \rangle$ ).

Conditional distributions will be denoted using a standard notation, e.g.,  $\pi(y|X)$ , or, in case that we consider conditioning by a specific value of variable  $X$  by  $\pi(y|X = \mathbf{a})$ . This must be strictly differentiated from the result of intervention, for which we use Pearl's *do* operator; i.e.,  $\pi(y|do(X = \mathbf{a}))$  denotes the effect on  $y$  caused by the intervention  $do(X = \mathbf{a})$ , which sets variable  $X$  to a value  $\mathbf{a} \in \langle X \rangle$ .

For a probability distribution  $\pi(x)$  its *marginal* distribution for  $y \subset x$  is denoted either by  $\pi(y)$ , or by  $\pi^{\downarrow y}$ . The latter notation will especially be used when marginalizing a distribution expressed in the form of a formula, or when referring to the value of probability distribution  $\pi$  for a specific state  $\mathbf{a}$  (combination of values of variables):  $\pi^{\downarrow y}(\mathbf{a})$ . In this paper we are interested also in an opposite operation called *extension*. By this term we understand any distribution  $\kappa$  defined for a superset of variables, i.e.,  $\kappa(z)$  for  $z \supset x$ , such that  $\kappa(x) = \pi(x)$ . The set of all extensions of distribution  $\pi(x)$  for variables  $z \supset x$  will be denoted by  $\Psi[\pi; z]$ .

Now, consider two distributions  $\pi(x)$  and  $\kappa(y)$ . Obviously, there exists their joint extension if and only if they are *consistent*, i.e., if  $\pi(x \cap y) = \kappa(x \cap y)$ . In case that they are not consistent then one can be interested in getting an extension of  $\pi$  containing from  $\kappa$  *as much information as possible*. Speaking more precisely, one can look for a distribution  $\mu(x \cup y)$  that is a *projection* of  $\kappa$  into the set of all extensions  $\Psi[\pi; x \cup y]$ :

$$\mu(x \cup y) = \arg \min_{\lambda \in \Psi[\pi; x \cup y]} Div(\lambda(y); \kappa(y)).$$

If the considered divergence is the Kullback-Leibler divergence<sup>1</sup>

$$Div(\lambda(y); \kappa(y)) = \begin{cases} \sum_{\mathbf{a} \in \langle y \rangle : \lambda^{\downarrow y}(\mathbf{a}) > 0} \lambda^{\downarrow y}(\mathbf{a}) \log \frac{\lambda^{\downarrow y}(\mathbf{a})}{\kappa(\mathbf{a})} & \text{if } \lambda^{\downarrow y} \ll \kappa, \\ +\infty & \text{otherwise,} \end{cases}$$

then Theorem 6.2 in [3] states that this type of projection can be got as a composition of  $\pi$  and  $\kappa$  that is defined by the formula

$$\mu(x \cup y) = \pi(x) \triangleright \kappa(y) = \begin{cases} \frac{\pi(x) \cdot \kappa(y)}{\kappa(x \cap y)} & \text{if } \pi(x \cap y) \ll \kappa(x \cap y), \\ \text{undefined} & \text{otherwise.} \end{cases} \quad (1)$$

Since the operator of composition is one of the key notions of this paper let us recall its most important properties proven in [3].

---

<sup>1</sup> $\lambda(y) \ll \kappa(y)$  denote that  $\kappa(y)$  *dominates*  $\lambda(y)$ , which holds (in the considered finite setting) when

$$\kappa(\mathbf{b}) = 0 \implies \lambda(\mathbf{b}) = 0$$

for all  $\mathbf{b} \in \langle y \rangle$ .

**Theorem 2.1.** Suppose  $\pi(x)$ ,  $\kappa(y)$  and  $\lambda(z)$  are probability distributions such that  $\kappa^{\downarrow x \cap y} \gg \pi^{\downarrow x \cap y}$ ,  $\lambda^{\downarrow(x \cup y) \cap z} \gg (\pi \triangleright \kappa)^{\downarrow(x \cup y) \cap z}$  and  $\lambda^{\downarrow x \cap z} \gg \pi^{\downarrow x \cap z}$ . Then the following statements hold:

1. (*Domain*):  $\pi \triangleright \kappa$  is a probability distribution for  $x \cup y$ .
2. (*Composition preserves first marginal*):  $(\pi \triangleright \kappa)^{\downarrow x} = \pi$ .
3. (*Reduction*): If  $y \subseteq x$  then,  $\pi \triangleright \kappa = \pi$ .
4. (*Non-commutativity*): In general,  $\pi \triangleright \kappa \neq \kappa \triangleright \pi$ .
5. (*Commutativity under consistency*):  $\pi$  and  $\kappa$  are consistent if and only if  $\pi \triangleright \kappa = \kappa \triangleright \pi$ .
6. (*Non-associativity*): In general,  $(\pi \triangleright \kappa) \triangleright \lambda \neq \pi \triangleright (\kappa \triangleright \lambda)$ .
7. (*Associativity under special condition I*): If  $x \supset (y \cap z)$  then,  $(\pi \triangleright \kappa) \triangleright \lambda = \pi \triangleright (\kappa \triangleright \lambda)$ .
8. (*Associativity under special condition II*): If  $y \supset (x \cap z)$  then,  $(\pi \triangleright \kappa) \triangleright \lambda = \pi \triangleright (\kappa \triangleright \lambda)$ .
9. (*Stepwise composition*): If  $(x \cap y) \subseteq w \subseteq y$  then,  $(\pi \triangleright \kappa^{\downarrow w}) \triangleright \kappa = \pi \triangleright \kappa$ .
10. (*Exchangeability*): If  $x \supset (y \cap z)$  then,  $(\pi \triangleright \kappa) \triangleright \lambda = (\pi \triangleright \lambda) \triangleright \kappa$ .
11. (*Simple marginalization*): If  $(x \cap y) \subseteq w \subseteq (x \cup y)$  then,  $(\pi \triangleright \kappa)^{\downarrow w} = \pi^{\downarrow x \cap w} \triangleright_{\kappa^{\downarrow y \cap w}}$ .

The reader certainly noticed that the assumptions on the dominance relations in the previous assertion guarantee that the compositions appearing in the individual properties are well defined. To make the following exposition more lucid, and to avoid repeating again and again the necessary assumptions on the dominance relation, let us make the following convention: Whenever the operator of composition appears in this paper, the respective relation of dominance is supposed to hold, so that all the compositions are always well defined. Notice that this can be trivially secured by supposing that all the distributions are strictly positive, but in this case we would lost a possibility to model logical (deterministic) dependence.

Recall that Property 6 from Theorem 2.1 states that the composition operator is generally not associative. This disadvantage is, in a way, compensated by the existence of an alternative operator, called *anticipating composition operator*, which is defined as follows. Consider probability distributions  $\pi(x)$  and  $\kappa(y)$  and another subset of variables  $z$ . Then,

$$\pi(x) \circledast_z \kappa(y) = (\pi(x) \cdot \kappa((z \setminus x) \cap y)) \triangleright \kappa(y). \quad (2)$$

Notice that this composition operator is parameterized by subset  $z$ . If  $(z \setminus x) \cap y = \emptyset$  then  $\pi \circledast_z \kappa = \pi \triangleright \kappa$ . An important property of this operator was proven in [3].

**Theorem 2.2.** For probability distributions  $\lambda(z)$ ,  $\pi(x)$  and  $\kappa(y)$  it holds that

$$(\lambda \triangleright \pi) \triangleright \kappa = \lambda \triangleright (\pi \oslash_z \kappa). \quad (3)$$

Adopting notation from [11] we will define a degenerate one-dimensional distribution  $\delta_{\mathbf{a}}(X)$ , which is a distribution of variable  $X$  achieving probability 1 for value  $X = \mathbf{a} \in \langle X \rangle$ , i.e.,

$$\delta_{\mathbf{a}}(X) = \begin{cases} 1 & \text{if } X = \mathbf{a}, \\ 0 & \text{otherwise.} \end{cases}$$

So, distribution  $\delta_{\mathbf{a}}(X)$  carries the sure information that  $X = \mathbf{a}$ . This provides a possibility to formulate the following assertion concerning conditional distributions.

**Theorem 2.3.** For probability distribution  $\pi(x)$  and  $x \supseteq y \cup \{X\}$ ,  $X \notin y$  it holds that

$$\pi(y|X = \mathbf{a}) = (\delta_{\mathbf{a}}(X) \triangleright \pi(x))^{\downarrow y}. \quad (4)$$

*Proof.* Denote  $z = y \cup \{X\}$ . Then, using Property 11 of Theorem 2.1, one can compute

$$(\delta_{\mathbf{a}}(X) \triangleright \pi(x))^{\downarrow y} = \left( (\delta_{\mathbf{a}}(X) \triangleright \pi(x))^{\downarrow z} \right)^{\downarrow y} = (\delta_{\mathbf{a}}(X) \triangleright \pi(z))^{\downarrow y}.$$

Property 1 of Theorem 2.1 says that  $\delta_{\mathbf{a}}(X) \triangleright \pi(z)$  is a probability distribution of variables  $z$ , and from Formula (1) defining the operator of composition it is clear that  $\delta_{\mathbf{a}}(X) \triangleright \pi(y, X = \mathbf{b}) = 0$  if  $\mathbf{b} \neq \mathbf{a}$ . In opposite case, i.e., if  $\mathbf{b} = \mathbf{a}$ , then

$$\delta_{\mathbf{a}}(X) \triangleright \pi(z) = \frac{\delta_{\mathbf{a}}(X) \cdot \pi(y, X = \mathbf{a})}{\pi(X = \mathbf{a})} = \pi(y|X = \mathbf{a}).$$

Therefore, marginalizing

$$(\delta_{\mathbf{a}}(X) \triangleright \pi(z))^{\downarrow y} = \sum_{\mathbf{b} \in \langle X \rangle} \delta_{\mathbf{a}}(X) \triangleright \pi(y, X = \mathbf{b}) = \pi(y|X = \mathbf{a}),$$

we get a distribution of variables  $y$ , which finishes the proof.  $\square$

### 3. CAUSAL MODELS

Pearl's causal models [7] are expressed in a form of Bayesian networks whose directed graphs define the corresponding causal relations: all arrows head from causes to consequences. In other words, relation of parents and children in the graph correspond to the relation of causes and consequences.

Conditioning in a causal model is just a conditioning in the corresponding Bayesian network, and the intervention in a causal model is defined as a conditioning in a modified Bayesian network; the modification consists in deleting all the edges heading to the node, in which the intervention is realized. This is because the intervention is performed in spite of any influence from other nodes (variables) of the considered causal model.

Let us illustrate the necessity of distinguishing between conditioning and intervention by a trivial example.

### 3.1. Example

Let us go back to our cement trading example. To dispatch unusually high number of orders we have to prolong loading shifts. To describe this relation formally consider two (for simplicity binary) variables describing whether we have got unusually many orders (variable  $Q$ ), and whether we have a prolonged loading shift (variable  $L$ ). Naturally, since we assume that great quantity of orders  $Q = \mathbf{y}$  results in a prolonged shift  $L = \mathbf{y}$ , there is an evident causal relation between these variables:  $Q$  is a cause for  $L$ . Clearly, these variables are mutually dependent, and therefore, denoting the respective probability distribution  $\pi(Q, L)$ , we quite naturally expect that  $\pi(Q = \mathbf{y}|L = \mathbf{y}) > \pi(Q = \mathbf{y})$  and  $\pi(L = \mathbf{y}|Q = \mathbf{y}) > \pi(L = \mathbf{y})$ .

The situation changes when, instead of conditioning, we consider intervention. Using Pearl's notation we denote by  $do(Q = \mathbf{y})$  the situation when somebody arranges that we get a great number of orders to dispatch. Analogously,  $do(L = \mathbf{y})$  denotes the situation when we decide to organize a long shift. In this case, it is natural to expect that receiving an enormous number of orders results in the necessity to hold a prolonged loading shift, but organizing a prolonged shift does not increase the number of received orders. Therefore while  $\pi(L = \mathbf{y}|do(Q = \mathbf{y})) > \pi(L = \mathbf{y})$ ,  $\pi(Q = \mathbf{y}|do(L = \mathbf{y})) = \pi(Q = \mathbf{y})$ .

Let us express this consideration in a formal model. Consider variable  $X$  and a set of variables  $\mathfrak{C}(X)$  that are *causes* of  $X$ . It means that the behavior of variable  $X$  is fully described by some  $\pi(x)$ , for  $x = \{X\} \cup \mathfrak{C}(X)$ . Using Properties 3 and 5 from Theorem 2.1 one can immediately see that  $\pi(x) = \pi(\mathfrak{C}(X)) \triangleright \pi(x)$ , which seems to be unnecessarily complex, but which may be, as we are now going to show, for causal models very useful.

Theorem 2.3 says that

$$\pi(\mathfrak{C}(X)|X = \mathbf{a}) = (\delta_{\mathbf{a}}(X) \triangleright \pi(x))^{\downarrow \mathfrak{C}(X)}, \quad (5)$$

and taking into consideration that  $\pi(x) = \pi(\mathfrak{C}(X)) \triangleright \pi(x)$  we get

$$\pi(\mathfrak{C}(X)|X = \mathbf{a}) = (\delta_{\mathbf{a}}(X) \triangleright (\pi(\mathfrak{C}(X)) \triangleright \pi(x)))^{\downarrow \mathfrak{C}(X)}.$$

Knowing that the operator of composition is not associative, we know that generally

$$(\delta_{\mathbf{a}}(X) \triangleright (\pi(\mathfrak{C}(X)) \triangleright \pi(x)))^{\downarrow \mathfrak{C}(X)} \neq ((\delta_{\mathbf{a}}(X) \triangleright \pi(\mathfrak{C}(X))) \triangleright \pi(x))^{\downarrow \mathfrak{C}(X)}.$$

The difference between these two expressions follows immediately from the application of Properties 3 and 11 (see Theorem 2.1) to the right hand part of the above inequality

$$\begin{aligned} ((\delta_{\mathbf{a}}(X) \triangleright \pi(\mathfrak{C}(X))) \triangleright \pi(x))^{\downarrow \mathfrak{C}(X)} &= (\delta_{\mathbf{a}}(X) \triangleright \pi(\mathfrak{C}(X)))^{\downarrow \mathfrak{C}(X)} \\ &= (\delta_{\mathbf{a}}(X)^{\downarrow \emptyset} \triangleright \pi(\mathfrak{C}(X))) = \pi(\mathfrak{C}(X)). \end{aligned}$$

So, computing  $((\delta_{\mathbf{a}}(X) \triangleright \pi(\mathfrak{C}(X))) \triangleright \pi(x))^{\downarrow \mathfrak{C}(X)}$  we get the marginal  $\pi(\mathfrak{C}(X))$ . In the next section we will see that it must be this way because, as we will show,

$$((\delta_{\mathbf{a}}(X) \triangleright \pi(\mathfrak{C}(X))) \triangleright \pi(x))^{\downarrow \mathfrak{C}(X)} = \pi(\mathfrak{C}(X)|do(X = \mathbf{a})),$$

and, as illustrated by the example above, the intervention that changes (fixes) the value of an effect variable does not influence behavior of its causes.

### 3.2. Causal compositional models

In this paper we express the considered causal models in a form of compositional models, which are known to be equivalent to Bayesian networks (for details see [3]).

Let us consider a set of variables  $w = \{X_1, X_2, \dots, X_n\}$ , and for each variable  $X_i \in w$  let  $\mathfrak{C}(X_i) \subset w$  be its causes. Here we consider only *Markovian* models [7], i. e., the models in which variables can be ordered (without loss of generality we assume it is the ordering  $X_1, X_2, \dots, X_n$ ) such that causes are always before their effects. So, we assume that

$$X_k \in \mathfrak{C}(X_i) \implies k < i,$$

which, as the reader certainly noticed, means that  $\mathfrak{C}(X_1) = \emptyset$ , and excludes feedback models from our consideration.

In keeping with the above notation, denote  $x_i = \mathfrak{C}(X_i) \cup \{X_i\}$ , and let  $\pi_i(x_i)$  denote the distribution describing local behavior of  $X_i$ . This means that we consider a causal model<sup>2</sup>

$$\kappa(X_1, X_2, \dots, X_n) = \pi_1(x_1) \triangleright \pi_2(x_2) \triangleright \dots \triangleright \pi_n(x_n). \quad (6)$$

In the next section we will show how to compute the result of an intervention to a variable  $X \in w$ . To prove the respective formula we will have to show that it realizes conditioning in a modified causal model, in a model in which  $\mathfrak{C}(X) = \emptyset$ . So, we want now to find a causal compositional model, which differs from the considered model given by Formula 6 only in the fact that for the new model  $\mathfrak{C}(X) = \emptyset$ . Since we assume that the original model is Markovian, it is evident that the new model is Markovian, too. Let us show that the new model is<sup>3</sup>

$$\kappa'(X_1, X_2, \dots, X_n) = \pi_i(X) \triangleright \pi_1(x_1) \triangleright \pi_2(x_2) \triangleright \dots \triangleright \pi_n(x_n), \quad (7)$$

where

$$i = \min\{k : X \in x_k\}$$

(which means that  $X = X_i$ ). If  $i = 1$  then validity of expression (7) is obvious (it follows directly from Property 3 from Theorem 2.1). If  $i = 2$ , then first applying Property 5 (keep in mind that  $\mathfrak{C}(X_1) = \emptyset$ , and therefore  $x_1 = \{X_1\}$ , and  $x_1 \cap \{X\} = \emptyset$ ) and then Property 3 one gets

$$\begin{aligned} \pi_2(X) \triangleright \pi_1(x_1) \triangleright \pi_2(x_2) \triangleright \dots \triangleright \pi_n(x_n) &= \pi_1(x_1) \triangleright \pi_2(X) \triangleright \pi_2(x_2) \triangleright \dots \triangleright \pi_n(x_n) \\ &= \pi_1(x_1) \triangleright \pi_2(X) \triangleright \pi_3(x_3) \triangleright \dots \triangleright \pi_n(x_n). \end{aligned}$$

For  $i > 2$  we have first to apply Property 5, then  $(n - 2)$  times Property 10, and finally

---

<sup>2</sup>To avoid too many brackets in formulae, we do not use them if the operator of composition is performed from left to right, i. e.,

$$\pi_1(x_1) \triangleright \pi_2(x_2) \triangleright \dots \triangleright \pi_n(x_n) = (\dots(\pi_1(x_1) \triangleright \pi_2(x_2)) \triangleright \dots \triangleright \pi_{n-1}(x_{n-1})) \triangleright \pi_n(x_n).$$

<sup>3</sup>It is obvious that, due to Property 3 of Theorem 2.1, the term  $\pi_i(x_i)$  can be deleted from right hand side of Formula (7). We prefer not doing it since it leads to simpler and more elegant formulae.

again Property 3:

$$\begin{aligned}
\pi_i(X) \triangleright \pi_1(x_1) \triangleright \pi_2(x_2) \triangleright \dots \triangleright \pi_n(x_n) &= \pi_1(x_1) \triangleright \pi_i(X) \triangleright \pi_2(x_2) \triangleright \dots \triangleright \pi_n(x_n) \\
&= \pi_1(x_1) \triangleright \pi_2(x_2) \triangleright \pi_i(X) \triangleright \pi_3(x_3) \triangleright \dots \triangleright \pi_n(x_n) \\
&= \dots = \pi_1(x_1) \triangleright \pi_2(x_2) \triangleright \dots \triangleright \pi_{i-1}(x_{i-1}) \triangleright \pi_i(X) \triangleright \pi_i(x_i) \triangleright \dots \triangleright \pi_n(x_n) \\
&= \pi_1(x_1) \triangleright \pi_2(x_2) \triangleright \dots \triangleright \pi_i(X) \triangleright \dots \triangleright \pi_n(x_n).
\end{aligned}$$

### 3.3. Intervention in causal compositional models

**Theorem 3.1.** For the causal compositional model  $\kappa$  given by Formula (6), and for arbitrary  $X \in x_1 \cup \dots \cup x_n$ ,  $\mathbf{a} \in \langle X \rangle$ ,  $y \subseteq x_1 \cup \dots \cup x_n \setminus \{X\}$ ,

$$\kappa(y|do(X = \mathbf{a})) = \left( \delta_{\mathbf{a}}(X) \triangleright \pi_1(x_1) \triangleright \pi_2(x_2) \triangleright \dots \triangleright \pi_n(x_n) \right)^{\downarrow y}. \quad (8)$$

**P r o o f.** Since we follow Pearl's idea, we define the intervention as a conditioning in the altered model, i.e.,  $\kappa(y|do(X = \mathbf{a})) = \kappa'(y|X = \mathbf{a})$ , where  $\kappa'$  is given by Formula (7). Using Theorem 2.3

$$\begin{aligned}
\kappa(y|do(X = \mathbf{a})) &= \kappa'(y|X = \mathbf{a}) \\
&= \left( \delta_{\mathbf{a}}(X) \triangleright (\pi_i(X) \triangleright \pi_1(x_1) \triangleright \pi_2(x_2) \triangleright \dots \triangleright \pi_n(x_n)) \right)^{\downarrow y}.
\end{aligned}$$

The expression  $\delta_{\mathbf{a}}(X) \triangleright (\pi_i(X) \triangleright \pi_1(x_1) \triangleright \pi_2(x_2) \triangleright \dots \triangleright \pi_n(x_n))$  can further be simplified using  $n$  times Property 8 and at last once Property 3

$$\begin{aligned}
\delta_{\mathbf{a}}(X) \triangleright &\left( \pi_i(X) \triangleright \pi_1(x_1) \triangleright \pi_2(x_2) \triangleright \dots \triangleright \pi_n(x_n) \right) \\
&= \delta_{\mathbf{a}}(X) \triangleright \left( \pi_i(X) \triangleright \pi_1(x_1) \triangleright \pi_2(x_2) \triangleright \dots \triangleright \pi_{n-1}(x_{n-1}) \right) \triangleright \pi_n(x_n) = \dots \\
&= \delta_{\mathbf{a}}(X) \triangleright \pi_i(X) \triangleright \pi_1(x_1) \triangleright \pi_2(x_2) \triangleright \dots \triangleright \pi_{n-1}(x_{n-1}) \triangleright \pi_n(x_n) \\
&= \delta_{\mathbf{a}}(X) \triangleright \pi_1(x_1) \triangleright \pi_2(x_2) \triangleright \dots \triangleright \pi_{n-1}(x_{n-1}) \triangleright \pi_n(x_n),
\end{aligned} \quad (9)$$

and therefore

$$\kappa(y|do(X = \mathbf{a})) = (\delta_{\mathbf{a}}(X) \triangleright \pi_1(x_1) \triangleright \pi_2(x_2) \triangleright \dots \triangleright \pi_{n-1}(x_{n-1}) \triangleright \pi_n(x_n))^{\downarrow y},$$

which was to be proven.  $\square$

Quite often we are interested in a effect of intervention  $do(X = \mathbf{a})$  to a single variable  $Y \in y$ . Though the focus of this paper is not oriented to finding optimum computational algorithms (for algorithmic solution of computational problems the author is referred to Malvestuto's papers [5, 6]), we want to show that principally the general expression (8) can be further transformed into simpler formulae. The level of simplification is dependent on the first appearance of  $Y$  in Formula (6). Therefore, we keep the notation

$$i = \min\{k : X \in x_k\},$$

and denote

$$j = \min\{k : Y \in x_k\}.$$

When modifying Formula (8) we will distinguish two situations according to the relation of values  $i$  and  $j$ : whether  $j < i$ , or  $j > i$  (obviously,  $i \neq j$ , because each  $x_i = \{X_i\} \cup \mathfrak{C}(X_i)$ ).<sup>4</sup>

First, assume  $j < i$ .

In this case applying Property 5 and afterwards several times (precisely  $(j - 1)$ -times) Property 10 we get

$$\begin{aligned} \delta_{\mathbf{a}}(X) \triangleright \pi_1(x_1) \triangleright \pi_2(x_2) \triangleright \dots \triangleright \pi_n(x_n) &= \pi_1(x_1) \triangleright \delta_{\mathbf{a}}(X) \triangleright \pi_2(x_2) \triangleright \dots \triangleright \pi_n(x_n) \\ &= \pi_1(x_1) \triangleright \pi_2(x_2) \triangleright \dots \triangleright \pi_j(x_j) \triangleright \delta_{\mathbf{a}}(X) \triangleright \pi_{j+1}(x_{j+1}) \triangleright \dots \triangleright \pi_i(x_i) \triangleright \dots \triangleright \pi_n(x_n), \end{aligned}$$

and therefore (using Property 2)

$$\begin{aligned} \kappa(Y|do(X = \mathbf{a})) &= \left( \delta_{\mathbf{a}}(X) \triangleright \pi_1(x_1) \triangleright \pi_2(x_2) \triangleright \dots \triangleright \pi_n(x_n) \right)^{\downarrow\{Y\}} \\ &= \left( \pi_1(x_1) \triangleright \pi_2(x_2) \triangleright \dots \triangleright \pi_j(x_j) \right)^{\downarrow\{Y\}} \\ &= \kappa(Y). \end{aligned} \quad (10)$$

Now, consider the other possibility  $j > i$ .

Using analogous reasoning as in the previous step we start computing

$$\begin{aligned} (\delta_{\mathbf{a}}(X) \triangleright \pi_1(x_1) \triangleright \dots \triangleright \pi_n(x_n))^{\downarrow x_i \cup \dots \cup x_j} \\ = (\pi_1(x_1) \triangleright \dots \triangleright \pi_{i-1}(x_{i-1}) \triangleright \delta_{\mathbf{a}}(X) \triangleright \pi_i(x_i) \triangleright \dots \triangleright \pi_j(x_j))^{\downarrow x_i \cup \dots \cup x_j}. \end{aligned}$$

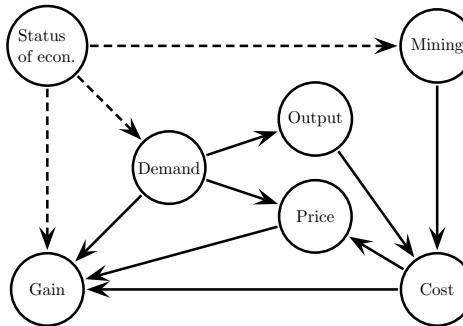
First notice that the term  $\pi_i(x_i)$  can be deleted because of Property 3. Further, denoting  $x_i \cup \dots \cup x_j = z$ , we can apply several times Property 11

$$\begin{aligned} &(\delta_{\mathbf{a}}(X) \triangleright \pi_1(x_1) \triangleright \dots \triangleright \pi_n(x_n))^{\downarrow z} \\ &= (\pi_1(x_1) \triangleright \dots \triangleright \pi_{i-1}(x_{i-1}) \triangleright \delta_{\mathbf{a}}(X) \triangleright \pi_{i+1}(x_{i+1}) \triangleright \dots \triangleright \pi_j(x_j))^{\downarrow z} \\ &= (\pi_1(x_1) \triangleright \dots \triangleright \pi_{i-1}(x_{i-1}) \triangleright \delta_{\mathbf{a}}(X) \triangleright \pi_{i+1}(x_{i+1}) \triangleright \dots \triangleright \pi_{j-1}(x_{j-1}))^{\downarrow z \setminus \{X_j\}} \triangleright \pi_j(x_j) \\ &= \dots \\ &= (\pi_1(x_1) \triangleright \dots \triangleright \pi_{i-1}(x_{i-1}))^{\downarrow z \setminus \{X_i, \dots, X_j\}} \triangleright \delta_{\mathbf{a}}(X) \triangleright \pi_{i+1}(x_{i+1}) \triangleright \dots \triangleright \pi_j(x_j), \end{aligned}$$

from which we can deduce a general formula

$$\kappa(Y|do(X = \mathbf{a})) = \left( (\pi_1(x_1) \triangleright \dots \triangleright \pi_{i-1}(x_{i-1}))^{\downarrow z \setminus \{X_i, \dots, X_j\}} \triangleright \delta_{\mathbf{a}}(X) \triangleright \right. \\ \left. \triangleright \pi_{i+1}(x_{i+1}) \triangleright \dots \triangleright \pi_j(x_j) \right)^{\downarrow\{Y\}}. \quad (11)$$

<sup>4</sup>The reader familiar with Bayesian network theory certainly noticed that in case that  $X$  and  $Y$  are independent, then their mutual placement in the sequence can be changed by selecting another, fully equivalent ordering of variables. This fact, naturally, plays an important role in algorithmic solutions of the computational problems, but this, as said at another place, is not the goal of this paper.



**Fig. 2.** Causal model: factors influencing the gain in cement trading.

The last question to be answered in this section is how to compute the effect of a multiple intervention. But it is an easy task. Consider a causal compositional model  $\kappa(w) = \pi_1(x_1) \triangleright \dots \triangleright \pi_n(x_n)$ , and (for simplicity) two variables  $V, X \in w$ . Let  $\mathbf{a} \in \langle X \rangle$ ,  $\mathbf{b} \in \langle V \rangle$ , and denote  $y = w \setminus \{V, X\}$ . Then

$$\begin{aligned} \kappa(y|do((V, X) = (\mathbf{b}, \mathbf{a})) &= \kappa(y|do(V = \mathbf{b}), do(X = \mathbf{a})) \\ &= (\delta_{\mathbf{b}}(V) \triangleright \delta_{\mathbf{a}}(X) \triangleright \pi_1(x_1) \triangleright \pi_2(x_2) \triangleright \dots \triangleright \pi_n(x_n))^{\downarrow y} \\ &= (\delta_{(\mathbf{b}, \mathbf{a})}(V, X) \triangleright \pi_1(x_1) \triangleright \pi_2(x_2) \triangleright \dots \triangleright \pi_n(x_n))^{\downarrow y}. \end{aligned}$$

#### 4. EXAMPLES

When evaluating the impact on variable  $Y$  of an intervention to variable  $X$  we usually need to take into account other related factors. And quite often these related factors cannot be observed (some authors call them *confounders* or *covariates*). In this context a natural question arises: Can we evaluate the effect of intervention even in the presence of hidden (unobservable) confounders? The answer is, in some cases, positive (see, e.g., [7, 10]) and in the rest of this paper we present simple examples illustrating how to use the formalism of causal compositional models for this purpose. To be able to describe all four examples just with one causal model, let us enrich the cement trade model from Figure 1 by four more variables as depicted in Figure 2. So, in all the following examples we consider the following causal compositional model

$$\begin{aligned} \kappa(C, D, G, M, O, P, S) = \pi_1(S) \triangleright \pi_2(D, S) \triangleright \pi_3(M, S) \triangleright \pi_4(D, O) \triangleright \\ \triangleright \pi_5(C, M, O) \triangleright \pi_6(C, D, P) \triangleright \pi_7(C, D, G, P, S). \end{aligned}$$

##### 4.1. Example 1

First, let us consider a simple situation when we intend to evaluate the effect of restriction of the production output  $O$  on the demand  $D$ , i.e., our aim is to compute  $\kappa(D|do(O = \mathbf{o}))$ . Following the recipe from Section 3 we realize that  $i = 4$  and  $j = 2$  and we can

directly use the Formula (10) as a result for  $j < i$  and obtain simple result

$$\kappa(D|do(O = \mathbf{o})) = (\pi_1(S) \triangleright \pi_2(D, S))^{\downarrow\{D\}}.$$

We can see that the result contains a marginal (of the original model) in variable  $D$  which can be estimated from data, and that the considered intervention has, in the considered model, no impact on variable  $D$ .

#### 4.2. Example 2

If the cement traders want to analyze the effect of restriction of the production output  $O$  on the price  $P$ , then, according to the Section 3, we get  $i = 4$  and  $j = 6$ , which leads to the application of Formula (11). So, in this case we arrive at

$$\begin{aligned} \kappa(P|do(O = \mathbf{o})) = & \left( (\pi_1(S) \triangleright \pi_2(D, S) \triangleright \pi_3(M, S))^{\downarrow\{D, M\}} \triangleright \delta_{\mathbf{o}}(O) \triangleright \right. \\ & \left. \triangleright \pi_5(C, M, O = \mathbf{o}) \triangleright \pi_6(C, D, P) \right)^{\downarrow\{P\}}. \end{aligned}$$

Again we succeeded in elimination of the hidden variable  $S$ , the marginal of original model  $(\pi_1(S) \triangleright \pi_2(D, S) \triangleright \pi_3(M, S))^{\downarrow\{D, M\}}$  is distribution in variables  $D$  and  $M$  and can be estimated from data. This is an example of situation fulfilling the back-door criterion (see Pearl [7]) and the result is the same as the one obtained using Pearl's back-door adjustment.

#### 4.3. Example 3

If evaluating  $\pi(G|do(P = \mathbf{p}))$ , it is apparent that we should proceed according to Formula (11), because in this case we get  $i = 6$  and  $j = 7$ . However, using this formula we cannot eliminate the hidden variable  $S$  because  $S$  is a direct cause of  $G$ . In this case we obtain an expression

$$\begin{aligned} \kappa(G|do(P = \mathbf{p})) &= \left( (\pi_1(S) \triangleright \pi_2(D, S) \triangleright \pi_3(M, S) \triangleright \pi_4(D, O) \triangleright \pi_5(C, M, O)) \right. \\ &\quad \left. \triangleright \delta_{\mathbf{p}}(P) \triangleright \pi_7(C, D, G, P, S) \right)^{\downarrow\{G\}}, \end{aligned}$$

which cannot be evaluated from observable data.

Nevertheless, several properties from Theorem 2.1 provide a possibility to eliminate the hidden variable and compute the effect of intervention even in this case. The elicitation of the causal effect is as follows (the labels above the signs of equation refer to

the properties of Theorem 2.1 used):

$$\begin{aligned}
\kappa(G|do(P = \mathbf{p})) &= \left( \delta_{\mathbf{p}}(P) \triangleright \pi_1(S) \triangleright \pi_2(D, S) \triangleright \pi_3(M, S) \triangleright \pi_4(D, O) \triangleright \pi_5(C, M, O) \triangleright \right. \\
&\quad \left. \triangleright \pi_6(C, D, P) \triangleright \pi_7(C, D, G, P, S) \right)^{\downarrow\{G\}} \\
&\stackrel{4 \times (8)}{=} \left( \delta_{\mathbf{p}}(P) \triangleright (\pi_1(S) \triangleright \pi_2(D, S) \triangleright \pi_3(M, S) \triangleright \pi_4(D, O) \triangleright \pi_5(C, M, O)) \triangleright \right. \\
&\quad \left. \triangleright \pi_6(C, D, P) \triangleright \pi_7(C, D, G, P, S) \right)^{\downarrow\{G\}} \\
&\stackrel{(9)}{=} \left( \delta_{\mathbf{p}}(P) \triangleright (\pi_1(S) \triangleright \pi_2(D, S) \triangleright \pi_3(M, S) \triangleright \pi_4(D, O) \triangleright \pi_5(C, M, O)) \right)^{\downarrow\{C,D\}} \triangleright \\
&\quad \triangleright (\pi_1(S) \triangleright \pi_2(D, S) \triangleright \pi_3(M, S) \triangleright \pi_4(D, O) \triangleright \pi_5(C, M, O)) \triangleright \\
&\quad \left. \triangleright \pi_6(C, D, P) \triangleright \pi_7(C, D, G, P, S) \right)^{\downarrow\{G\}} \\
&\stackrel{(7)}{=} \left( \delta_{\mathbf{p}}(P) \triangleright (\pi_1(S) \triangleright \dots \triangleright \pi_5(C, M, O)) \right)^{\downarrow\{C,D\}} \triangleright (\pi_1(S) \triangleright \pi_2(D, S) \triangleright \pi_3(M, S) \triangleright \right. \\
&\quad \left. \triangleright \pi_4(D, O) \triangleright \pi_5(C, M, O) \triangleright \pi_6(C, D, P) \right) \triangleright \pi_7(C, D, G, P, S) \Big)^{\downarrow\{G\}} \\
&\stackrel{(8)}{=} \left( \delta_{\mathbf{p}}(P) \triangleright (\pi_1(S) \triangleright \dots \triangleright \pi_5(C, M, O)) \right)^{\downarrow\{C,D\}} \triangleright (\pi_1(S) \triangleright \pi_2(D, S) \triangleright \pi_3(M, S) \triangleright \right. \\
&\quad \left. \triangleright \pi_4(D, O) \triangleright \pi_5(C, M, O) \triangleright \pi_6(C, D, P) \triangleright \pi_7(C, D, G, P, S) \right)^{\downarrow\{G\}} \\
&\stackrel{(11)}{=} \left( \delta_{\mathbf{p}}(P) \triangleright \kappa(C, D, G, M, O, P, S)^{\downarrow\{C,D\}} \triangleright \kappa(C, D, G, M, O, P, S)^{\downarrow\{C,D,G\}} \right)^{\downarrow\{G\}}.
\end{aligned}$$

So, even in this case we obtain a result containing the marginals of the model that can be estimated from the data.

#### 4.4. Example 4

The last example describes an (admittedly rather artificial) situation, when the government is considering to run huge road infrastructure project and minister of industry wants to know the impact of increased demand for cement  $D$  on the producer's cost  $C$ , i.e., we want to evaluate

$$\begin{aligned}
\kappa(C|do(D = \mathbf{d})) &= \left( \delta_{\mathbf{d}}(D) \triangleright \pi_1(S) \triangleright \pi_2(D, S) \triangleright \pi_3(M, S) \triangleright \pi_4(D, O) \triangleright \right. \\
&\quad \left. \triangleright \pi_5(C, M, O) \triangleright \pi_6(C, D, P) \triangleright \pi_7(C, D, G, P, S) \right)^{\downarrow\{C\}}.
\end{aligned}$$

Also in this case we need to employ particular properties from Theorem 2.1 because, again, hidden variable  $S$  is the parent of  $D$ . First of all we shall simplify the model because variables  $P$  and  $G$  cannot influence the demand. Let us eliminate  $G$  and then  $P$

(and hence also both  $\pi_6$  and  $\pi_7$ ) using the pair of Properties 11 and 3 twice. After this, we get the following expression, which is further modified:

$$\begin{aligned}
& \kappa(C|do(D = \mathbf{d})) \\
&= \left( \delta_{\mathbf{d}}(D) \triangleright \pi_1(S) \triangleright \pi_2(D, S) \triangleright \pi_3(M, S) \triangleright \pi_4(D, O) \triangleright \pi_5(C, M, O) \right)^{\downarrow\{C\}} \\
&\stackrel{(3)}{=} \left( \delta_{\mathbf{d}}(D) \triangleright \pi_1(S) \triangleright \pi_3(M, S) \triangleright \pi_4(D, O) \triangleright \pi_5(C, M, O) \right)^{\downarrow\{C\}} \\
&\stackrel{(8)}{=} \left( \delta_{\mathbf{d}}(D) \triangleright (\pi_1(S) \triangleright \pi_3(M, S)) \triangleright \pi_4(D, O) \triangleright \pi_5(C, M, O) \right)^{\downarrow\{C\}} \\
&\stackrel{(10)}{=} \left( \delta_{\mathbf{d}}(D) \triangleright \pi_4(D, O) \triangleright (\pi_1(S) \triangleright \pi_3(M, S)) \triangleright \pi_5(C, M, O) \right)^{\downarrow\{C\}} \\
&= \left( \delta_{\mathbf{d}}(D) \triangleright \pi_4(D, O) \triangleright \left( (\pi_1(S) \triangleright \pi_3(M, S)) \oslash_{\{D, O\}} \pi_5(C, M, O) \right) \right)^{\downarrow\{C\}},
\end{aligned}$$

where the property number is again indicated upon the equality sign. The last modification was performed according to Theorem 2.2. In order to evaluate the causal effect we need to eliminate the hidden variable  $S$ , thus, let us focus our attention to the corresponding subexpression

$$\begin{aligned}
& \left( (\pi_1(S) \triangleright \pi_3(M, S)) \oslash_{\{D, O\}} \pi_5(C, M, O) \right)^{\downarrow\{C, M, O\}} \\
&= \left( (\pi_1(S) \triangleright \pi_3(M, S)) \cdot \pi_5(O) \triangleright \pi_5(C, M, O) \right)^{\downarrow\{C, M, O\}} \\
&\stackrel{(3)}{=} \left( (\pi_1(S) \triangleright \pi_2(S) \triangleright \pi_3(M, S)) \cdot \pi_5(O) \triangleright \pi_5(C, M, O) \right)^{\downarrow\{C, M, O\}} \\
&\stackrel{(11)}{=} \left( (\pi_1(S) \triangleright \pi_2(D, S) \triangleright \pi_3(M, S))^{\downarrow\{M, S\}} \cdot \pi_5(O) \triangleright \pi_5(C, M, O) \right)^{\downarrow\{C, M, O\}} \\
&\stackrel{(11)}{=} \left( (\pi_1(S) \triangleright \pi_2(D, S) \triangleright \pi_3(M, S)) \cdot \pi_5(O) \triangleright \pi_5(C, M, O) \right)^{\downarrow\{C, M, O\}}.
\end{aligned}$$

Notice that in the last formula, the expression  $(\pi_1(S) \triangleright \pi_2(D, S) \triangleright \pi_3(M, S))$  is a three-dimensional marginal of  $\kappa$ , i.e.  $\kappa(D, M, S)$ . It contains hidden variable  $S$  but having the data, we can easily estimate its two-dimensional marginal

$$\kappa(D, M) = (\pi_1(S) \triangleright \pi_2(D, S) \triangleright \pi_3(M, S))^{\downarrow\{D, M\}},$$

which enables us to eliminate variable  $S$ . So we get

$$\begin{aligned}
& \left( (\pi_1(S) \triangleright \pi_2(D, S) \triangleright \pi_3(M, S)) \cdot \pi_5(O) \triangleright \pi_5(C, M, O) \right)^{\downarrow\{C, M, O\}} \\
&\stackrel{(11)}{=} \left( (\pi_1(S) \triangleright \pi_2(D, S) \triangleright \pi_3(M, S))^{\downarrow\{D, M\}} \cdot \pi_5(O) \triangleright \pi_5(C, M, O) \right)^{\downarrow\{C, M, O\}} \\
&= \left( (\pi_1(S) \triangleright \pi_2(D, S) \triangleright \pi_3(M, S))^{\downarrow\{D, M\}} \oslash_{\{O\}} \pi_5(C, M, O) \right)^{\downarrow\{C, M, O\}}.
\end{aligned}$$

All these computations show that we can evaluate the effect of the considered intervention according to the formula

$$\begin{aligned} \kappa(C|do(D = \mathbf{d})) &= \left( \delta_{\mathbf{d}}(D) \triangleright \pi_4(D, O) \triangleright \right. \\ &\quad \left. \triangleright \left( (\pi_1(S) \triangleright \pi_2(D, S) \triangleright \pi_3(M, S))^{\downarrow\{D, M\}} \odot_{\{O\}} \pi_5(C, M, O) \right)^{\downarrow\{C, M, O\}} \right)^{\downarrow\{C\}}, \end{aligned} \quad (12)$$

where the marginal  $\kappa(D, H) = (\pi_1(S) \triangleright \pi_2(D, S) \triangleright \pi_3(M, S))^{\downarrow\{D, M\}}$  can be estimated from data.

Let us remark that the expression

$$\left( (\pi_1(S) \triangleright \pi_2(D, S) \triangleright \pi_3(M, S))^{\downarrow\{D, M\}} \odot_{\{O\}} \pi_5(C, M, O) \right)^{\downarrow\{C, M, O\}}$$

contains variable  $D$  which is, however, marginalized out in this expression. Here we employ Pearl's idea of extension used in his front-door adjustment [7]. The reader can realize that it is the way how to take into account the mutual dependence of variables  $C, D, O$  (notice that it plays the same role that is realized by inheritance of parents in Shachters edge reversal rule [9]).

## 5. CONCLUSIONS

We have presented an alternative algebraical approach to model the effect of an external intervention. To achieve this goal the proposed methodology employs causal compositional models and its properties. In Section 3 we derived general formulae for the evaluation of causal effect in models where latent variable does not appear among the parents of effect variable. Studied possibilities are illustrated by four examples showing both the cases when general method is applicable and when it fails. The presented examples show that the causal effect of an intervention can be evaluated even in some cases of models containing latent variables. Notice that the structure of expression (12) goes beyond the usual compositional sequences; it corresponds to more general non-sequential models whose structures are studied by Malvestuto in [5].

## ACKNOWLEDGEMENT

The authors want to express their great acknowledgement to an anonymous referee whose constructive comments helped to substantially increase the legibility of the paper. The research was supported by GA ČR under grant no. 15-00215S.

(Received March 1, 2015)

## REFERENCES

---

- [1] A. Detwarasiti and R. D. Shachter: Influence diagrams for team decision analysis. *Decision Analysis* 2 (2005), 4, 207–228. DOI:10.1287/deca.1050.0047

- [2] Y. Hagmayer, S. Sloman, D. Lagnado, and M.R. Waldmann: Causal reasoning through intervention. In: Causal Learning: Psychology, Philosophy, and Computation (A. Gopnik and L. Schulz, eds.), Oxford University Press 2007, pp. 86–101. DOI:10.1093/acprof:oso/9780195176803.003.0007
- [3] R. Jiroušek: Foundations of compositional model theory. *Int. J. Gen. Syst.* 40 (2011), 6, 623–678. DOI:10.1080/03081079.2011.562627
- [4] R. Jiroušek: On causal compositional models: Simple examples. In: Proc. 15th Int. Conf. on Inf. Processing and Management of Uncertainty – Part I. Springer 2014, pp. 517–526. DOI:10.1007/978-3-319-08795-5\_53
- [5] F. M. Malvestuto: Equivalence of compositional expressions and independence relations in compositional models. *Kybernetika* 50 (2014), 3, 322–362. DOI:10.14736/kyb-2014-3-0322
- [6] F. M. Malvestuto: Marginalization in models generated by compositional expressions. To appear in *Kybernetika* 51 (2015), 4.
- [7] J. Pearl: Causality: Models, Reasoning, and Inference. Cambridge University Press, NY 2009. DOI:10.1017/cbo9780511803161
- [8] M. Ryall and A. Bramson: Inference and Intervention: Causal Models for Business Analysis. Routledge, NY 2013. DOI:10.4324/9780203076835
- [9] R. Shachter: Evaluating influence diagrams. *Oper. Res.* 34 (1986), 6, 871–882. DOI:10.1287/opre.34.6.871
- [10] P. Spirtes, C. Glymour, and R. Scheines: Causation, Prediction and Search. Springer Lecture Notes in Statistics, New York 1993. DOI:10.1007/978-1-4612-2748-9
- [11] R. R. Tucci: Introduction to Judea Pearl’s Do-Calculus. arXiv:1305.5506v1 [cs.AI] (2013).

*Vladislav Bína, Faculty of Management in Jindřichův Hradec, University of Economics in Prague, Jarošovská 1117/II, 37701 Jindřichův Hradec. Czech Republic.  
e-mail: bina@fm.vse.cz*

*Radim Jiroušek, Faculty of Management in Jindřichův Hradec, University of Economics in Prague and Institute of Information Theory and Automation — the Czech Academy of Sciences, Pod Vodárenskou věží 4, 182 08 Praha 8. Czech Republic.  
e-mail: radim@utia.cas.cz*