Using Tensor-Type Formalism in Causal Networks

Claudiu Pozna[a, b], Radu-Emil Precup[c, d], Aron Ballagi[a]

[a] Department of Automation, Széchenyi István University, Egyetem tér 1, 9026 Győr, Hungary

[b] Department of Automation and Information Technology, Transilvania University of Brasov, Str. Mihai Viteazu 5, Corp V, et. 3, 500174 Brasov, Romania

[c] Department of Automation and Applied Informatics, Politehnica University of Timisoara, Bd. V. Parvan 2, 300223 Timisoara, Romania

[d] Center for Fundamental and Advanced Technical Research, Romanian Academy – Timisoara Branch, Bd. Mihai Viteazu 24, 300223 Timisoara, Romania

E-mail: cp@unitbv.ro, radu.precup@upt.ro, ballagi@sze.hu

Abstract: The causal network is a possible description of complex phenomena, and several domains, for example, Machine Learning, Social Science, and Artificial Intelligence. Although a successful solution is referred to in this paper, the field inherently faces challenges. Among these, the work identified that the formalism used is time-consuming and difficult to understand. Consequently, the approach proposed in this paper consists in transcribing this formalism in a tensor form. This goal is accomplished in three steps: first common tensor formulas are proposed for direct and inverse models; second these formulas are adapted for the network primitives; in the end the primitive and consequently the formula composition is analysed. To facilitate the understanding of the proposed formalism, the paper describes several examples. This paper is dedicated to Prof. Imre J. Rudas, to celebrate his 75th anniversary.

Keywords: Causal networks; Formalism; Tensors

1 Introduction

Frequently, the description of a phenomenon is done with a causal network model. These networks are complex constructions of events, between which causal relationships can be established [1] [2]. Causality is illustrated using a directed graph attached to the network. On the graph, the connection between two events A and B is in the form of an arrow (A→ B) which implies the information that event A is the cause of event B [3] [4].

The network contains three types of elements: roots, source events, those that are not caused; intermediate, those that are caused by other events and, in turn, cause other events; respectively the leaves that are caused but do not cause.

We will call the dependence (model) A→B a direct model. Contrary the connection B←A, which represents the questioning of the causes that produced a known effect, we will call it the inverse model. By themselves, the two dependencies can be deterministic or stochastic. We will focus on those of the stochastic type that can be customized in the deterministic version. Starting from here, we can imagine several variants, among which we mention the probabilistic and the possibilistic (fuzzy). Of these, we will deal with the first.

Like any source of knowledge, the causal network provides answers to questions (which will be symbolized by Q_i). They refer to probability distributions of the variables of an event, eventually under the conditions of certain observations.

Causal networks are applied in various domains: epidemiology [5] for identifying the causes of diseases and assessing the effectiveness of interventions; machine learning, where causal inference [6] is increasingly used to make predictions and decisions in complex systems; social sciences [7] for understanding the impact of policies, interventions, and socio-economic factors; Artificial Intelligence (AI) as a generalization of machine learning [8] [9], which benefits from causal reasoning for explainability and robust decision-making.

Future directions in the field include developing methods for causal discovery, improving causal inference algorithms, and enhancing the integration of causality into AI systems.

The causal network challenges include unobserved confounders [2], selection bias, and data limitations. We also include here the challenge of the used mathematical formalism which is time-consuming and makes understanding difficult. Our proposal starts from this point. It consists of a tensor-like construction of the model associated with the network, this assembly combines two types of dependencies: the direct and the inverse models.

This paper is dedicated to Prof. Imre J. Rudas, on his 75th anniversary. The first two authors of this chapter are grateful to Prof. Rudas for enabling their association, which led to a fruitful cooperation for more than 15 years.

The present paper is organized into four parts: the introduction where we define the paper objective; the section dedicated to the proposed formalism presentation, where the tensor-type general formula is adapted for the network primitives and the possibility of composition is analyzed; Section 3 contains an example of using the proposed formalism; conclusions and future work intention end the paper.

2 The Proposed Formalism

The paper proposes a tensor formalism that allows easy manipulation of causal models. To achieve this desire, the strategy used is the identification of primitives, which through composition generate any causal network; the analysis of these primitives, i.e., obtaining the two types of models for each primitive; finally, the description of the synthesis method (primitives composition) that leads to the solution of the possible questions.

Some preliminary remarks regarding the notations used are necessary:

- An event A is characterized by the variable a which can have the values $a_{1,\ldots,n}$;
- The probability that the event A has the value a_i is denoted $P(A=a_i)$ or more simply $P(a_i)$;
- The probability distribution for the event A is denoted by a vector $P(A)=\left[P(a_1) \quad \ldots \quad P(a_n)\right]^T$, where $\sum_{i=1}^{n} P(a_i)=1$;
- The conditional probability distribution of the event B when we observe the variable $a=a_i$ is the vector $P(B \mid a_i)=\left[P(b_1 \mid a_i) \quad \ldots \quad P(b_m \mid a_i)\right]^T$, where:
- We call a closed event that event on which an observation was made directly or indirectly, for example, $w=w_i$ or $P(B \mid w_i)$ otherwise the event is called open.

Using the marginalization formula, respectively the one related to the probability of conjugate events (b_j and $a_{1,\ldots,n}$), the following representation of the direct causal model A→B is proposed:

$$P(b_j)=\sum_{i=1}^{n} P(b_j, a_i)=\left[P(b_j \mid a_1) \quad \ldots \quad P(b_j \mid a_n)\right]\cdot\left[P(a_1) \quad \ldots \quad P(a_n)\right]^T$$
$$=\left[P(b_j \mid a_1) \quad \ldots \quad P(b_j \mid a_n)\right]\cdot P(A)$$

$$P(B)=\left[P(b_1) \quad \ldots \quad P(b_m)\right]^T=\begin{bmatrix} P(b_1 \mid a_1) & \ldots & P(b_1 \mid a_n) \\ . & \ldots & . \\ P(b_m \mid a_1) & \ldots & P(b_m \mid a_n) \end{bmatrix}\cdot\begin{bmatrix} P(a_1) \\ . \\ P(a_n) \end{bmatrix}={}_{\mathbf{A}}^{\mathbf{B}}\mathbf{P}\cdot P(A)$$

$$\Rightarrow \tag{1}$$

$$P(B)={}_{\mathbf{A}}^{\mathbf{B}}\mathbf{P}\cdot P(A)$$

where: $\,^{\mathbf{B}}_{\mathbf{A}}\mathbf{P} = \begin{bmatrix} P(b_1 \mid a_1) & \dots & P(b_1 \mid a_n) \\ . & \dots & . \\ P(b_m \mid a_1) & \dots & P(b_m \mid a_n) \end{bmatrix}$ is the direct model A→B; in $\,^{\mathbf{B}}_{\mathbf{A}}\mathbf{P}$ the lower index refers to the cause and the upper index refers to the effect.

Bayes' theorem is used for the inverse model:

$$P(A \mid b_j) = \begin{bmatrix} P(a_1 \mid b_j) & \dots & P(a_n \mid b_j) \end{bmatrix}^T = \begin{bmatrix} P(a_1)\dfrac{P(b_j \mid a_1)}{P(b_j)} & \dots & P(a_n)\dfrac{P(b_j \mid a_n)}{P(b_j)} \end{bmatrix}^T =$$

$$\begin{bmatrix} P(a_1)\dfrac{P(b_j \mid a_1)}{\sum_{i=1}^{n} P(a_i)P(b_j \mid a_i)} & \dots & P(a_n)\dfrac{P(b_j \mid a_n)}{\sum_{i=1}^{n} P(a_i)P(b_j \mid a_i)} \end{bmatrix}^T$$

$$P(A \mid b_j) = \frac{\,^{\mathbf{b_j}}_{\mathbf{A}}\mathbf{P}^T \otimes P(A)}{\,^{\mathbf{b_j}}_{\mathbf{A}}\mathbf{P} \cdot P(A)} \tag{2}$$

where: $\,^{\mathbf{b_j}}_{\mathbf{A}}\mathbf{P}^T \otimes P(A) = \begin{bmatrix} P(b_j \mid a_1) & \dots & P(b_j \mid a_n) \end{bmatrix}^T \otimes \begin{bmatrix} P(a_1) & \dots & P(a_n) \end{bmatrix}^T = \begin{bmatrix} P(a_1)P(b_j \mid a_1) & \dots & P(a_n)P(b_j \mid a_n) \end{bmatrix}^T$; $\otimes$ is the symbol of the element-by-element product; $\,^{\mathbf{b_j}}_{\mathbf{A}}\mathbf{P} = \begin{bmatrix} P(b_j \mid a_1) & \dots & P(b_j \mid a_n) \end{bmatrix} = \,^{\mathbf{B}}_{\mathbf{A}}\mathbf{P}(j,:)$ is the row j of the direct model matrix.

Consequently, the two direct and inverse models are expressed as:

$$\begin{cases} A \to B : P(B) = \,^{\mathbf{B}}_{\mathbf{A}}\mathbf{P} \cdot P(A) \\ B = b_j \leftarrow A : P(A \mid b_j) = \dfrac{\,^{b_j}_{\mathbf{A}}\mathbf{P}^T \otimes P(A)}{\,^{b_j}_{\mathbf{A}}\mathbf{P} \cdot P(A)} \end{cases} \tag{3}$$

2.1 The Primitives

According to the mentioned work strategy after obtaining the general formulae of the direct and the inverse model, a special emphasis will be given to adapting them for the network structure. For this reason, three primitives have been identified:

- the causal chain consisting of two or more consecutive events;
- causal fork, where an event causes two, or more, different events;
- the causal collimator where an event is caused by two, or more, different events.

The three primitives will be analysed from the point of view of the direct and inverse models. The causal chain (Figure 1) is a simple network composed of two or more elements connected in series in a cause-effect sequence.

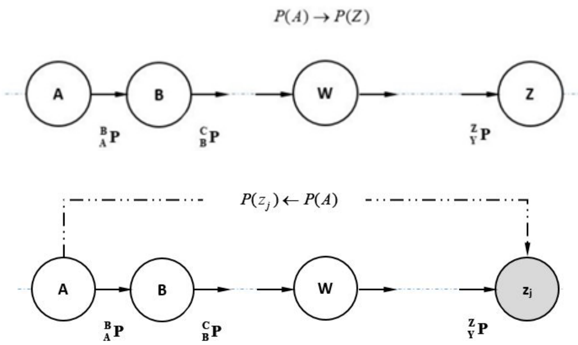


Figure 1

Causal chain: a) the direct model; b) the inverse model

The following relationships are written for the direct model:

$$\begin{aligned}
&P(B)={}^{\mathbf{B}}_{\mathbf{A}}\mathbf{P}\cdot P(A)\\
&P(C)={}^{\mathbf{C}}_{\mathbf{B}}\mathbf{P}\,{}^{\mathbf{B}}_{\mathbf{A}}\mathbf{P}\cdot P(A)\\
&\ldots\\
&P(Z)={}^{\mathbf{Z}}_{\mathbf{Y}}\mathbf{P}\ldots{}^{\mathbf{C}}_{\mathbf{B}}\mathbf{P}\,{}^{\mathbf{B}}_{\mathbf{A}}\mathbf{P}\cdot P(A)\\
&\Rightarrow {}^{\mathbf{Z}}_{\mathbf{A}}\mathbf{P}=\prod_{W=B}^{W=Z}{}^{\mathbf{W}}_{\mathbf{A}}\mathbf{P}
\end{aligned}\tag{4}$$

For the inverse model (Figure 1 b)), observing the variable z_j leads to:

$$P(A\mid z_j)=\frac{{}^{z_j}_{\mathbf{A}}\mathbf{P}^T\otimes P(A)}{{}^{z_j}_{\mathbf{A}}\mathbf{P}\cdot P(A)}\tag{5}$$

It is interesting to analyse the effect of the observation of the variable w_j on the direct model (Figure 2). According to the (accepted) Markovian hypothesis, this observation separates the causal chain into two segments in which models of the type $w_j \leftarrow A$ and $w_j \rightarrow Z$ make sense. The events A and Z become independent.

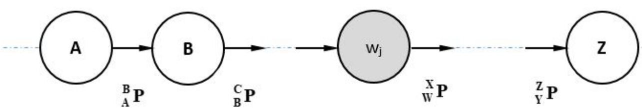


Figure 2

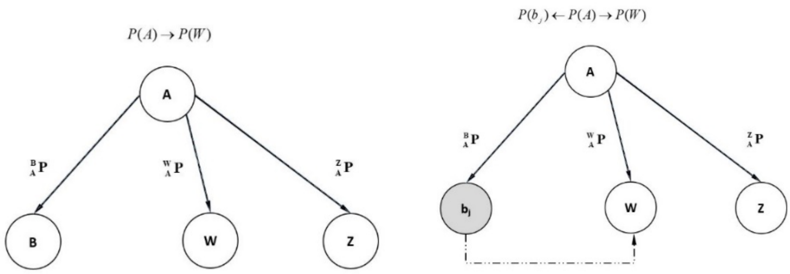
The influence of observation

$$P(A \mid w_j) = \frac{{}^{w_j}_{\mathbf{A}}\mathbf{P}^T \otimes P(A)}{{}^{w_j}_{\mathbf{A}}\mathbf{P} \cdot P(A)} \tag{6}$$

$$P(Z) = {}^{\mathbf{Z}}_{\mathbf{W}}\mathbf{P} \cdot \begin{bmatrix} 0 & . & 1_j & . & 0 \end{bmatrix}^T \tag{7}$$

Consequently, the observation w_j closes the events A, B, ..., W.

A causal fork (Figure 3 a)) is a simple network composed of three or more elements connected so that the first is the cause of the other ones.



a) b)

Figure 3

Causal fork: a) direct model; b) inverse model

The direct model A→{B, C, …, Z} is (Figure 3 a)):

$$P(W) = {}^{\mathbf{W}}_{\mathbf{A}}\mathbf{P} \cdot P(A) \tag{8}$$

where: $W \in \{B, C, ..., Z\}$.

The inverse model can refer to the event A ($b_j \leftarrow A$) or the events B, C, ..., Z ($b_j \leftarrow A \to W$) (also illustrated in Figure 3 b)), it is observed that, in this case, {B, C, ... Z} are not independent events.

$$P(A \mid b_j) = \frac{{}^{b_j}_{\mathbf{A}}\mathbf{P}^T \otimes P(A)}{{}^{b_j}_{\mathbf{A}}\mathbf{P} \cdot P(A)} \tag{9}$$

$$P(W \mid b_j) = {}^{\mathbf{W}}_{\mathbf{A}}\mathbf{P} \cdot \frac{{}^{b_j}_{\mathbf{A}}\mathbf{P}^T \otimes P(A)}{{}^{b_j}_{\mathbf{A}}\mathbf{P} \cdot P(A)} \tag{10}$$

Observing the event a_j, breaks the connection between the events, in this case {B, C,..., Z} become independent events (Figure 4).

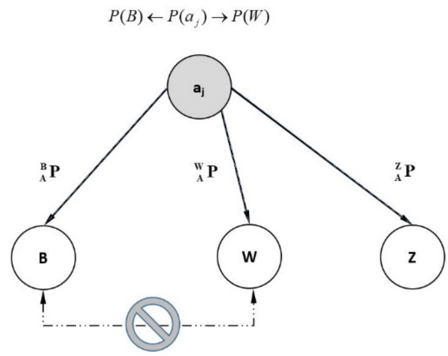


Figure 4
Causal fork: a) direct model; b) inverse model

The collimator (Figure 5 a)) is a simple network composed of three or more elements connected in such a way that one element is caused by the rest of the elements.

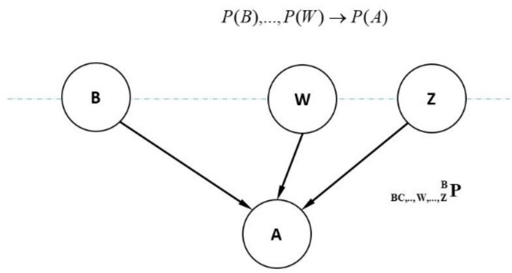


Figure 5
The collimator direct model

$$P(A) = \sum_{jb,jc,...,jz} P(A, b_{jb}, c_{jc}, ..., z_{jz})$$

$$= \sum_{jb,jc,...,jz} P(z_{jz}),...,P(c_{jc}) \cdot P(b_{jb}) P(A \mid b_{jb}, c_{jc}, ..., z_{jz}) =$$

$$= \sum_{jz} P(z_{jz}),...,\sum_{jc} P(c_{jc}) \sum_{jb} P(b_{jb}) P(A \mid b_{jb}, c_{jc}, ..., z_{jz}) =$$

$$= {}^{\mathbf{A}}_{\mathbf{BC,...,Z}}\mathbf{P} \cdot P(B)P(C),...,P(Z)$$

$$P(A) = {}^{\mathbf{A}}_{\mathbf{BC,...,Z}}\mathbf{P} \cdot P(B)P(C),...,P(Z) \tag{11}$$

where: $P(B)_{(n_b \times 1)},...,P(Z)_{(n_z \times 1)}$ are the probability distributions of the events B,…, Z; ${}^{\mathbf{A}}_{\mathbf{BC,...,Z}}\mathbf{P}_{(n_a \times n_b \times,...,n_z)}$ the model of the collimator i.e. a tensor.

For the direct model, the events B, C, …, and Z are independent. The inverse model assumes the observation of event A. In this case events B, C, ..., and Z become dependent:

$$P(W \mid a_j) = \frac{{}^{a_j}_{\mathbf{BC}}\mathbf{P}^T P(B),...,P(Z) \otimes P(W)}{{}^{a_j}_{\mathbf{BC}}\mathbf{P} P(B),...,P(Z) \cdot P(W)} \quad (12)$$

For example, the direct model of a collimator composed of 3 elements {B, C}→A is

$$\begin{aligned} P(A) &= \sum_{jb,jc} P(A, b_{jb}, c_{jc}) = \sum_{jb,jc} P(c_{jc}) \cdot P(b_{jb}) P(A \mid b_{jb}, c_{jc}) = \\ &= \sum_{jc} P(c_{jc}) \sum_{jb} P(b_{jb}) P(A \mid b_{jb}, c_{jc}) = {}^{\mathbf{A}}_{\mathbf{BC}}\mathbf{P} P(B) P(C) \end{aligned} \quad (13)$$

Figure 6 illustrates a graphical representation of the tensor multiplication.

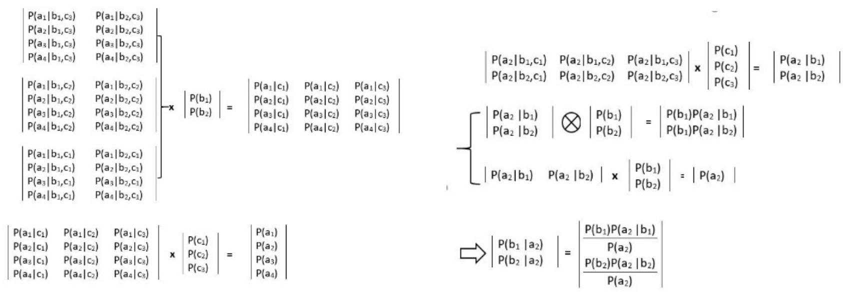


Figure 6

Using the tensor formalism for the direct and inverse model of a collimator-type structure

$$P(B \mid a_j) = \frac{{}^{a_j}_{\mathbf{BC}}\mathbf{P}^T P(C) \otimes P(B)}{{}^{a_j}_{\mathbf{BC}}\mathbf{P} P(C) \cdot P(B)} \quad (14)$$

The composition rule, here successive multiplication, of the $n_a \times,...,\times n_z$ size tensor with $n_{a,b,...,z} \times 1$ size vectors can be written as follows:

$$\begin{aligned} & n_a \times \underline{n_b}\{n_c, n_d, ..., n_z\} \circ \underline{n_b} \times 1 \rightarrow n_a \times n_c \{n_d, n_e, ..., n_z\} \\ & n_a \times \underline{n_c}\{n_d, n_e, ..., n_z\} \circ \underline{n_c} \times 1 \rightarrow n_a \times n_d \{n_e, n_f, ..., n_z\} \\ & ... \\ & n_a \times \underline{n_z} \circ \underline{n_z} \times 1 \rightarrow n_a \times 1 \end{aligned} \quad (15)$$

where: $\circ$ is the composition operator symbol (here multiplication); $\sum_{i=1}^{n_a} P(a_i \mid b_{jb}, c_{jc}, ..., z_{jz}) = 1$, i.e. the sum of the elements on the columns of the matrices that compose the tensor.

2.2 Composing the Primitives in Paths

A path in the causal network is a successive causal stepping from one node to another, on the vortex that connects them. This operation can have several purposes:

- one may ask which is the effect of the probability distribution of the root nodes on all child nodes, and especially on the (last) leaf nodes;
- one may also ask on the influence of the observations (of the value of the variables) against the probability distribution from the rest of the nodes.

The knowledge required for the answers was included in the study of the three primitives. We propose a synthesis of the primitives in (active) network paths. More precisely, we are interested in when the variable A can influence the variable Y, knowing that we have observations on the variables in the set O.

To proceed with the computation of the path-related models, we begin by defining the incidence matrix of the causal network. This matrix highlights the links of the events, their type (root, leaf, intermediate), the primitives of which the respective element is a part, as well as the possible observations.

The incident matrix $I : \{A, B, ..., Z\} \times \{A, B, ..., Z\} \to \mathbf{Z}$ is defined actually as a function in terms of:

$$I(k,h) = \begin{cases} 1 & if \quad \exists h \to k \\ -j & if \quad \exists k_j \leftarrow h \\ 0 & else \end{cases}; \; I(k,k) = \begin{cases} -j & if \quad \exists k_j \\ \\ 0 & else \end{cases} \tag{16}$$

where: A, B, …, Z are the element of the events set, **Z** is the set of integer numbers; $\exists h \to k$ means that exist a causal connection from H to K; $\exists k_j$ means that the event K was observed, $k=k_j$.

Consequently:

- If $I(k,h) \neq 1 \forall h$ the event K is a root;
- If $I(k,h) \neq 1 \forall k$ the event H is a leaf;
- If $I(k, a, ..., z) = 1 \quad for \quad more\_then\_1$ the event K is a collimator $K \leftarrow C = \{W \mid I(k,w) = 1\}$, C is the set of the cause events, K is the effect event;

- If $I(k,k) = -j$ means $k = k_j$;
- If $I(a,...,z,h) = 1$ *for more_then_1* the event H is a fork $H \to F = \{W \mid I(w,h) = 1\}$, F is the set of the effects, H is the cause event.

Using the knowledge about event dependency the following observations can be made about the possible paths:

- If $I(k,h) = 1$, and H is not a fork or a collimator and:
  - If $I(k,k) = 0$: then the path $H \to K$ is possible and it is described by the direct model;
  - If $I(k,k) = -j$: then the path $k_j \leftarrow H$ is possible and it is described by the inverse model, the observation will close the event H.
- If $I(k,h) = 1$, and $H \to \{W \mid I(w,h) = 1\}$ ($U, V, K,... \in F$) is a fork and:
  - If $I(h,h) = -j$: then the path $u_k \leftarrow H \to V$ is not possible
  - If $I(h,h) = 0$: then the path $u_k \leftarrow H \to V$ is possible by the inverse and direct model
  - If: then the path $H \to K$ is possible by using the direct model;
  - If $I(k,k) = -j$: then the path $k_j \leftarrow H$ is possible by using the inverse model.
- If $I(k,h) = 1$, and $K \leftarrow \{W \mid I(k,w) = 1\}$ ($U, V \in C$) is a collimator and:
  - If $I(k,k) = -j$ then the path $U \to k_j \leftarrow V$ is possible;
  - If $I(m,m) = -j$ and M is a descendant of K (K is closed by M) then the path $U \to m_j \leftarrow V$ is possible;
  - If $I(k,k) = 0$ then the path is not possible.

Table 1 summarizes the preceding information. The first column of the table refers to the three identified primitives: chain, fork, and collimator. The second column specifies the nature of the observations and their belonging to the path variables. It is specified that *desc* O is an attribute that specifies the descending variables (effects) of the variables included in the set O. The third column systematizes the verdict regarding the existence of (active) routes.

Table 1
Possible paths

The connection type	W	Observations
$A \rightarrow W \rightarrow Z$	$W \in \mathbf{O}$	No path, observation of W leads to independence of Z relative to A
	$W \notin \mathbf{O}$	There is a path from A to Z, a direct model
$w_j \leftarrow A$	$W \in \mathbf{O}$	The path from W to A exists, the reverse pattern
$B \leftarrow a_j \rightarrow Z$	$A \in \mathbf{O}$	There is no path from B to Z, observation of A leads to the independence of B relative to Z
$B \leftarrow A \rightarrow Z$	$A \notin \mathbf{O}$	There is a path from A to Z, a direct model
$b_j \leftarrow A \rightarrow Z$	$A \notin \mathbf{O}$	There is a path from B to Z, a reverse model
$B \rightarrow A \leftarrow Z$	$A \notin \mathbf{O}$	There is a path from B, Z to A direct model
$B \rightarrow A \leftarrow Z$	$A \notin \mathbf{O}$ and desc $A \notin \mathbf{O}$	No path from B to Z
$B \rightarrow a_j \leftarrow Z$	$A \in \mathbf{O}$ or desc $A \in \mathbf{O}$	There is a path from B to Z inverse model

3 Example

To facilitate the understanding of the proposed formalism, it is exemplified in this section in terms of a well-known case [10]. The probabilistic variables and models are contained in Table 2. The scenario concerns a vehicle in the vicinity of a grassy field and a person interacting with this environment.

Table 2
The direct models of the network

Variable	**Value**	**Model (ad-hoc)**
Rain **R:**	r_1 – yes; r_2 – no	$P(R) = [0.6, 0.4]^T$
Car wash **W:**	w_1 – yes; s_2 – no	$P(W) = [0.2, 0.8]^T$
Wet Grass **G:**	g_1 – yes; g_2 – no	${}_{\mathbf{RW}}^{\mathbf{G}}\mathbf{P}$: w_1 $\begin{bmatrix} r_1 w_1 & r_2 w_1 \end{bmatrix}$ $\begin{bmatrix} g_1 \\ g_2 \end{bmatrix}$ $\begin{bmatrix} 0.9 & 0.7 \\ 0.1 & 0.3 \end{bmatrix}$ w_2 $\begin{bmatrix} r_1 w_2 & r_2 w_2 \end{bmatrix}$ $\begin{bmatrix} g_1 \\ g_2 \end{bmatrix}$ $\begin{bmatrix} 0.6 & 0.2 \\ 0.4 & 0.8 \end{bmatrix}$

Slip **S:**	s_1 – yes; s_2 – no	${}^{S}_{G}\mathbf{P}: \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} \begin{matrix} [g_1 & g_2] \\ \begin{bmatrix} 0.7 & 0.4 \\ 0.3 & 0.6 \end{bmatrix} \end{matrix}$
Grass growing **C :**	c_1 – yes; c_2 – no	${}^{C}_{G}\mathbf{P}: \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \begin{matrix} [g_1 & g_2] \\ \begin{bmatrix} 0.6 & 0.2 \\ 0.4 & 0.8 \end{bmatrix} \end{matrix}$
Trimming **T:**	t_1 – yes; t_2 – no	${}^{T}_{C}\mathbf{P}: \begin{bmatrix} t_1 \\ t_2 \end{bmatrix} \begin{matrix} [c_1 & c_2] \\ \begin{bmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{bmatrix} \end{matrix}$

The modelled phenomenon (Figure 7) can be described with the help of the following events: the vehicle may be (recently) washed (W) or it may have rained (R); both events are causes of the wet grass (G) event; in turn it produces the slip (S) event (the person slides on the lawn) and the grass growth (C) event; lawn growth is the cause of the Trimming (T) operation.

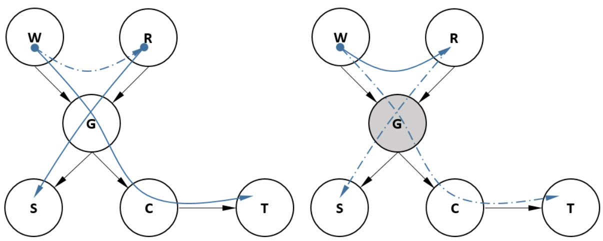
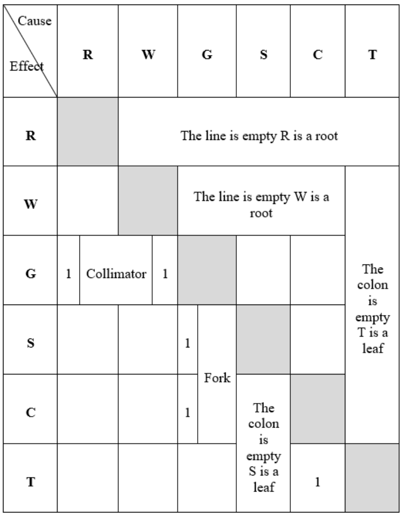


Figure 7
The example causal diagram

In Figure 8, the mentioned causal network is illustrated, and the six probabilistic variables can be observed as well as the connections between them. Also, from the multitude of possibilities, two variants appear, the one in which the variable G is not observed, respectively the one in which it is observed (g_1). The construction of the incidence matrix is illustrated in Figure 8.

In Figure 8, some possible routes are marked with a continuous line and some impossible ones with a broken line. Admissible (active) paths mean the possibility to answer certain questions. The proposed formalism was transposed into the Matlab programming environment and the results obtained after running the program are grouped in Table 3 and organized as simulation results.



$$\mathbf{I}=\begin{bmatrix}0&0&0&0&0&0\\0&0&0&0&0&0\\1&1&0&0&0&0\\0&0&1&0&0&0\\0&0&1&0&0&0\\0&0&0&0&1&0\end{bmatrix}$$

$$\mathbf{I}=\begin{bmatrix}0&0&0&0&0&0\\0&0&0&0&0&0\\1&1&-1&0&0&0\\0&0&1&0&0&0\\0&0&1&0&0&0\\0&0&0&0&1&0\end{bmatrix}(g_1)$$

Figure 8

The network incident matrix

Table 3

Simulations results

Question	Solution
Q1: What is the probability of cutting the grass? Here, the root nodes can be assigned the probability distribution from the table or we can assign a deterministic model for one of the values (w_1=1)	$P(T)={}_{\mathbf{C}}^{\mathbf{T}}\mathbf{P}\cdot P(C)$ $P(T)={}_{\mathbf{C}}^{\mathbf{T}}\mathbf{P}\cdot{}_{\mathbf{G}}^{\mathbf{C}}\mathbf{P}\cdot P(G)$ $P(T)={}_{\mathbf{C}}^{\mathbf{T}}\mathbf{P}\cdot{}_{\mathbf{G}}^{\mathbf{C}}\mathbf{P}\cdot{}_{\mathbf{RW}}^{\mathbf{G}}\mathbf{P}P(R)P(W)$ $P(T)=\begin{bmatrix}0.4845\\0.5155\end{bmatrix}$
Q2: What is the probability of mowing the grass when I notice that I am slipping? Here, the observed value is a_1	$P(T\mid s_1)={}_{\mathbf{C}}^{\mathbf{T}}\mathbf{P}\cdot P(C\mid s_1)$ $P(T\mid s_1)={}_{\mathbf{C}}^{\mathbf{T}}\mathbf{P}\cdot{}_{\mathbf{G}}^{\mathbf{C}}\mathbf{P}\cdot P(G\mid s_1)$ $P(T\mid s_1)={}_{\mathbf{C}}^{\mathbf{T}}\mathbf{P}\cdot{}_{G}^{C}\mathbf{P}\cdot\dfrac{{}_{\mathbf{G}}^{\mathbf{a_1}}\mathbf{P}^{T}\otimes P(G)}{{}_{\mathbf{G}}^{\mathbf{a_1}}\mathbf{P}\cdot P(G)}$ $P(T\mid s_1)={}_{\mathbf{C}}^{\mathbf{T}}\mathbf{P}\cdot{}_{\mathbf{G}}^{\mathbf{C}}\mathbf{P}\cdot\dfrac{{}_{\mathbf{G}}^{\mathbf{a_1}}\mathbf{P}^{T}\otimes{}_{\mathbf{RW}}^{\mathbf{G}}\mathbf{P}P(R)P(W)}{{}_{\mathbf{G}}^{\mathbf{a_1}}\mathbf{P}\cdot{}_{\mathbf{RW}}^{\mathbf{G}}\mathbf{P}P(R)P(W)}$ $P(T\mid a_1)=\begin{bmatrix}0.5223\\0.4777\end{bmatrix}$

Q3: What is the probability of having washed the car when I noticed that the lawn was wet? Here, the observed value is g_1	$P(S \mid g_1) = \dfrac{{}^{\mathbf{g_1}}_{\mathbf{R}W}\mathbf{P}^T P(R) \otimes P(W)}{{}^{\mathbf{g_1}}_{\mathbf{R}W}\mathbf{P}P(R) \cdot P(W)}$ $P(S \mid g_1) = \begin{bmatrix} 0.3178 \\ 0.6822 \end{bmatrix}$

This example can be easily deployed in other real-world applications. Such applications of authors' formalism could be selected in several fields including path planning [11] [12], electric vehicles [13] [14], haptic interfaces [15], Kalman filters [16], fuzzy control [17-21], manufacturing processes [22] [23], human well-being and resilience [24], pose estimation for robotic navigation [25], observation process modelling [26], and piezoelectric active laminated shells [27]. Various classical and metaheuristic optimization algorithms can be involved in these applications for performance improvement and also reducing the heuristics in setting the free parameters in the design approaches; they include cellular genetic algorithms [28], tabu search based on quantum computing [29], Bacterial Foraging Optimization Algorithms [30] [31], Clonal Selection Algorithms [31], Grey Wolf Optimizers [32], Particle Swarm Optimization Algorithms [33], Slime Mold Algorithms [23], classical iterative numerical optimization algorithms [34] [35], and metaheuristic algorithms with parameter adaptation [36].

Conclusions

The paper proposed a tensor-like formalism for causal network modelling. This alternative formalism uses matrix manipulation. Because the formulas that used summation on several variables are replaced by matrix operation the main benefit is the easier understanding of the model construction, the transformations can be observed for general cases in (1-3) and are afterward refined for different primitives (4,5,8,10,11,12) and possible compositions. Although deriving the model calculation which uses matrix computation will decrease the computing time.

The paper was structured in four parts. The introduction defines the paper's objective and presents its structure. The second section reveals the proposed formalism: initially, the general formula for the direct and inverse models are derived into a tensor form; this step is followed by the network primitive identification, which includes the general formula particularisation and in the end the composition of primitive and formulas identification. The paper ends with an example of network modelling by tensors.

Future research will be focused on making as transparent as possible authors' formalism and software focusing on real-world validation that includes the examples specified in the final part of previous section.

Acknowledgement

This work was supported by a grant of the Ministry of Research, Innovation and Digitization, CNCS/CCCDI - UEFISCDI, project number ERANET-ENUAC-e-MATS, within PNCDI IV, a grant of the Romanian Ministry of Research, Innovation and Digitization, CNCS/CCCDI – UEFISCDI, project number PN-III-P2-2.1-PED-2019-4366 (431PED) within PNCDI III, and the Széchenyi István University, Győr, Hungary.

References

[1] J. Pearl: Causality: Models, Reasoning, and Inference. Cambridge: Cambridge University Press, 2009

[2] D. Koller, N. Friedman: Probabilistic Graphical Models: Principles and Techniques. Cambridge, MA, London: The MIT Press, 2009

[3] P. Spirtes, C. Glymour, R. Scheines: Causation, Prediction, and Search, 2nd Ed. Cambridge, MA, London: The MIT Press, 2001

[4] F. V. Jensen: Bayesian Networks and Decision Graphs. New York, NY: Springer, 2007

[5] D. J. Spiegelhalter, J. P. Myles, D. R. Jones, K. R. Abrams: Bayesian methods in health technology assessment: a review, Health Technology Assessment, Vol. 4, No. 38, 2000, pp. 1-130

[6] J. Peters, D. Janzing, B. Schölkopf: Elements of Causal Inference: Foundations and Learning Algorithms. Cambridge, MA, London: The MIT Press, 2017

[7] G. W. Imbens, D. B. Rubin: Causal Inference for Statistics, Social, and Biomedical Sciences. Cambridge: Cambridge University Press, 2015

[8] M. A. Hernán, J. M. Robins: Causal Inference: What If. Boca Raton, FL: CRC Press, Taylor & Francis, 2023

[9] J. Pearl, M. Glymour, N. P. Jewell: Causal Inference in Statistics: A Primer. Chichester: John Wiley & Sons, 2016

[10] R. E. Neapolitan: Learning Bayesian Networks. Upper Saddle River, NJ: Prentice Hall, 2004

[11] G. Klančar, S. Blažič: Optimal constant acceleration motion primitives, IEEE Transactions on Vehicular Technology, Vol. 68, No. 9, 2019, pp. 8502-8511

[12] J. Vaščák, J. Hvizdoš, M. Puheim: Agent-based cloud computing systems for traffic management, in Proceedings of 2016 International Conference on Intelligent Networking and Collaborative Systems, Ostrava, Czech Republic, 2016, pp. 73-79

[13] Z. C. Johanyák: A modified particle swarm optimization algorithm for the optimization of a fuzzy classification subsystem in a series hybrid electric vehicle, Technicki Vjesnik - Technical Gazette, Vol. 24, No. 2, 2017, pp. 295-301

[14] A. Lucchini, S. Formentin, M. Corno, D. Piga, S. M. Savaresi: Torque vectoring for high-performance electric vehicles: a data-driven MPC approach, IEEE Control Systems Letters, Vol. 4, No. 3, 2020, pp. 725-730

[15] N. Ando, P. Korondi, H. Hashimoto: Networked telemicromanipulation systems "Haptic Loupe", IEEE Transactions on Industrial Electronics, Vol. 51, No. 6, 2004, pp. 1259-1271

[16] J. Kuti, I. J. Rudas, H.-J. Gao, P. Galambos: Computationally relaxed unscented Kalman filter, IEEE Transactions on Cybernetics, Vol. 53, No. 3, 2023, pp. 1557-1565

[17] R.-E. Precup, S. Preitl: Popov-type stability analysis method for fuzzy control systems, in Proceedings of Fifth European Congress on Intelligent Technologies and Soft Computing, Aachen, Germany, 1997, Vol. 2, pp. 1306-1310

[18] R. E. Haber, J. R. Alique, A. Alique, J. Hernández, R. Uribe-Etxebarria: Embedded fuzzy-control system for machining processes: Results of a case study, Computers in Industry, Vol. 50, No. 3, 2003, pp. 353-366

[19] M. L. Tomescu, S. Preitl, R.-E. Precup, J. K. Tar: Stability analysis method for fuzzy control systems dedicated controlling nonlinear processes, Acta Polytechnica Hungarica, Vol. 4, No. 3, 2007, pp. 127-141

[20] R.-E. Precup, T. Haidegger, S. Preitl, B. Benyó, A. S. Paul, L. Kovács: Fuzzy control solution for telesurgical applications, Applied and Computational Mathematics, Vol. 11, No. 3, 2012, pp. 378-397

[21] J.-B. Qiu, T. Wang, K.-K. Sun, I. J. Rudas, H.-J. Gao: Disturbance observer-based adaptive fuzzy control for strict-feedback nonlinear systems with finite-time prescribed performance, IEEE Transactions on Fuzzy Systems, Vol. 30, No. 4, 2022, pp. 1175-1184

[22] V. M. Nejković, M. S. Milićević, G. Janaćković, M. Grozdanović: Application of fuzzy analytic hierarchy process to inductive steel tube welding, Romanian Journal of Information Science and Technology, Vol. 25, No. 1, 2022, pp. 3-19

[23] R.-E. Precup, R.-C. David, R.-C. Roman, A.-I. Szedlak-Stinean, E. M. Petriu: Optimal tuning of interval type-2 fuzzy controllers for nonlinear servo systems using slime mould algorithm, International Journal of Systems Science, Vol. 54, No. 15, 2023, pp. 2941-2956

[24] F. G. Filip: Automation and computers and their contribution to human well-being and resilience, Studies in Informatics and Control, Vol. 30, No. 4, 2021, pp. 5-18

[25] C. Pozna, R.-E. Precup, P. Földesi: A novel pose estimation algorithm for robotic navigation, Robotics and Autonomous Systems, Vol. 63, 2014, pp. 10-21

[26] C. Pozna, R.-E. Precup: Aspects concerning the observation process modelling in the framework of cognition processes, Acta Polytechnica Hungarica, Vol. 9, No. 1, 2012, pp. 203-223

[27] P. Milić, D. Marinković, S. Klinge, Ž. Ćojbašić: Reissner-Mindlin based isogeometric finite element formulation for piezoelectric active laminated shells, Tehnički Vjesnik, Vol. 30, No. 2, 2023, pp. 416-425

[28] E. Osaba, J. Del Ser, A. D. Martinez, J. L. Lobo, F. Herrera: AT-MFCGA: An adaptive transfer-guided multifactorial cellular genetic algorithm for evolutionary multitasking, Information Sciences, Vol. 570, 2021, pp. 577-598

[29] E. Osaba, E. Villar-Rodriguez, I. Oregi, A. Moreno-Fernandez-de-Leceta: Hybrid quantum computing-tabu search algorithm for partitioning problems: preliminary study on the traveling salesman problem, in Proceedings of 2021 IEEE Congress on Evolutionary Computation, Kraków, Poland, 2021, pp. 351-358

[30] S. I. Bejinariu, H. Costin: Medical image registration by means of a bio-inspired optimization strategy, Computer Science Journal of Moldova, Vol. 20, No. 2, 2012, pp. 178-202

[31] S. I. Bejinariu, H. Costin, F. Rotaru, C. Niţă, R. Luca, C. Lazăr: Parallel processing and bio-inspired computing for biomedical image registration, Computer Science Journal of Moldova, Vol. 22, No. 2, 2014, pp. 253-277

[32] I. A. Zamfirache, R.-E. Precup, R.-C. Roman, E. M. Petriu: Neural network-based control using actor-critic reinforcement learning and grey wolf optimizer with experimental servo system validation, Expert Systems with Applications, Vol. 225, 2023, paper 120112

[33] D. Sánchez, P. Melin, O. Castillo: Comparison of particle swarm optimization variants with fuzzy dynamic parameter adaptation for modular granular neural networks for human recognition, Journal of Intelligent and Fuzzy Systems, Vol. 38, No. 3, 2020, pp. 3229-3252

[34] D. Singh, A. Shukla: Manifold optimization with MMSE hybrid precoder for Mm-Wave massive MIMO communication, Romanian Journal of Information Science and Technology, Vol. 25, No. 1, 2022, pp. 36-46

[35] S. Preitl, R.-E. Precup, J. Fodor, B. Bede: Iterative feedback tuning in fuzzy control systems. Theory and applications, Acta Polytechnica Hungarica, Vol. 3, No. 3, 2006, pp. 81-96

[36] O. Castillo, P. Melin, E. Ontiveros, C. Peraza, P. Ochoa, F. Valdez, J. Soria: A high-speed interval type 2 fuzzy system approach for dynamic parameter adaptation in metaheuristics, Engineering Applications of Artificial Intelligence, Vol. 85, 2019, pp. 666-680