A Direct Propagation Method in Singly Connected Causal Belief Networks with Conditional Distributions for all Causes

Oumaima Boussarsar, Imen Boukhris, and Zied Elouedi

LARODEC, Université de Tunis, Institut Supérieur de Gestion de Tunis, Tunisia
oumaima.boussarsar@hotmail.fr, imen.boukhris@hotmail.com,
zied.elouedi@gmx.fr

Abstract. Existing algorithms of propagation in belief networks deal with inference of observations when conditional distributions are initially defined per edge. The aim of this paper is to propose a direct method of causal inference of both observations and interventions on the causal belief networks quantified with the belief function theory where conditional beliefs are defined for all parents without having to transform the network into a junction tree. We explain how it is still possible to use the disjunctive rule of combination DRC and the generalized Bayesian theorem GBT to perform this propagation.

Keywords: Belief function theory, propagation, interventions, causal belief networks.

1 Introduction

Causality plays an important role in many fields, from physics to medicine to artificial intelligence. Indeed, causal knowledge simplifies decision-making. Interventions [6] are very useful for identifying causal relations. These latter are exterior manipulations that force target variables to have specific values. However, an observation is seeing and monitoring phenomena happening by themselves without any manipulation on the system.

The belief function theory is adequate to formalize imperfect causal knowledge that agents usually possess especially cases of ignorance. Accordingly, a graphical structure allows to simply represent and reason from such causal knowledge. Causal belief networks [3] are compact and flexible graphical representations where arcs are interpreted as causal links. On these networks, we can compute the effects of observations and also those of external actions.

In existing algorithms of propagation in belief networks either they are associational networks [1,11], or causal networks [4], the uncertainty is not modeled by a conditional mass function between a node and all its parents as for Bayesian networks, but as a set of local conditional distributions for a node and each of its parents. In the case where the expert gives conditional distributions defined for all parents, we must necessarily transform the network that is already simply

G.A. Aranda-Corral et al. (Eds.): AISC 2014, LNAI 8884, pp. 64–75, 2014.

connected (i.e., there are no two nodes that can be connected by more than one path) into a joint tree while the latter is usually used to transform multiply connected networks (i.e., arbitrary network structures) to a tree structure. Nodes in this tree are sets of variables called clusters. The propagation algorithm based on junction trees is expensive. Indeed, it depends on clusters' size.

In this paper, we propose a causal propagation method that performs directly on the initial causal belief network in the case where the conditional distributions are defined for all parents allowing to compute the effects of observations and also those of external actions. Our proposed algorithms are based on the two rules proposed by Smets [10] namely the disjunctive rule of combination (DRC) and the generalized Bayesian theorem (GBT). Moreover, we explain how these operators can be used on a set of variables.

The rest of the paper is organized as follows: in Section 2, we provide a brief background on the belief function theory. In Section 3, we recall causal belief networks. In Section 4, we explain how it is still possible to use the DRC and the GBT to perform this propagation. Inference in the presence of observations and interventions using mutilated and augmented graphs where conditional distributions are defined for all parents is described in Section 5. Section 6 concludes the paper.

2 Belief Function Theory

2.1 Definition

The theory of belief functions [8] is useful for representing uncertain knowledge. Let Θ be a finite non empty set including all the elementary events related to a given problem. These events are assumed to be exhaustive and mutually exclusive. Such Θ is called the frame of discernment. Beliefs are expressed on subsets belonging to the powerset of Θ denoted 2^{Θ}. The basic belief assignment (*bba*), denoted by m^{Θ} or m, is a mapping from 2^{Θ} to [0,1] such that: $\sum_{A \subseteq \Theta} m(A) = 1$. For each subset A of Θ, m(A) is called the basic belief mass (*bbm*). It represents the part of belief exactly committed to the event A of Θ. Subsets of Θ such that $m(\mathrm{A}) > 0$ are called focal elements. A *bba* is said to be certain if the whole mass is allocated to a unique singleton of Θ and Bayesian when all focal elements are singletons. If the *bba* has Θ as unique focal element, it is called vacuous and it represents the case of total ignorance.
The plausibility function *pl* quantified the maximum amount of belief that could be given to a subset A of Θ. It computes the total of masses compatible with A.

$$pl : 2^{\Theta} \rightarrow [0,1] \text{ such that:}$$

$$pl(A) = \sum_{A \cap C \neq \emptyset} m(C) \quad (1)$$

The basic belief assignment can be recovered from the plausibility function as follows:

$$m(A) = \sum_{C \subseteq A} (-1)^{|A-C+1|} pl(\bar{C}) \quad (2)$$

2.2 Basic Operations

Two *bba*s m_1 and m_2 provided by two distinct and independent sources, may be aggregated using Dempster's rule of combination, denoted by $\oplus$, as follows:

$$m_1 \oplus m_2(A) = K \cdot \sum_{B \cap C = A} m_1(B) m_2(C), \forall B, C \subseteq \Theta \tag{3}$$

where $K^{-1} = 1 - \sum_{B \cap C = \emptyset} m_1(B) m_2(C)$.

Smets [9] qualified Dempster's rule of conditioning as one of the natural ingredients and the center of the transferable belief model. Upon the arrival of a new information B, the initial knowledge encoded with a mass value, m(A), is revised using Dempster's rule of conditioning. $m(A|B)$ denotes the degree of belief of A in the context where B holds. It is defined as:

$$m(A|B) = \begin{cases} K. \sum_{C \subseteq \bar{B}} m(A \cup C) & \text{if} \quad A \subseteq B, A \neq \emptyset \\ 0 \quad \text{if} \quad A \nsubseteq B \end{cases} \tag{4}$$

where K^{-1}=$1 - m(\emptyset)$.

2.3 Multi-variable Operations

When we model aspects of the real world, the *bba*s induced from experts are defined on different frames of discernment. We recall in what follows useful multi-variables operations. Let us consider in what follows, a first frame Θ and a second frame Ω. A vacuous extension is changing the referential by adding new variables. Thus, a mass function m^{Θ} defined on Θ will be extended to $\Theta \times \Omega$ as follows:

$$m^{\Theta \uparrow \Theta \Omega}(C) = \begin{cases} m^{\Theta}(A) & \text{if} \quad C = A \times \Omega \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

Given a mass distribution defined on the product space $\Theta \times \Omega$, marginalization corresponds to mapping over a subset of the product space by dropping the extra coordinates. The new belief defined on Θ, $m^{\Theta \Omega \downarrow \Theta}$ is obtained by:

$$m^{\Theta \Omega \downarrow \Theta} = \sum_{C \subseteq \Theta \times \Omega, C^{\downarrow \Theta} = A} m^{\Theta \Omega}(C), A \subseteq \Theta \tag{6}$$

Smets [10] has generalized the Bayesian theorem within the transferable belief model framework known as the Generalized Bayesian Theorem (GBT). Let us consider $pl^{\Omega}(c|a_i)$ and $a_i \in a$ where $a \subseteq \Theta$ and $c \subseteq \Omega$. The a posteriori plausibility distribution $pl^{\Theta}(a|c)$ is defined as follows:

$$pl^{\Theta}(a|c) = 1 - \prod_{a_i \in a} (1 - pl^{C}(c|a_i)) \tag{7}$$

The function that is the dual of GBT is the disjunctive rule of combination (DRC). Let us consider $pl^{\Omega}(c|a_i)$ and $a_i \in a$ where $a \subseteq \Theta$ and $c \subseteq \Omega$. The plausibility distribution $pl(c|a)$ is defined as follows:

$$pl^{\Theta}(c|a) = pl^{\Theta}(a|c) = 1 - \prod_{a_i \in a} (1 - pl^{C}(c|a_i)) \tag{8}$$

3 Causal Belief Networks

Belief networks [1,3,11] are simple and efficient tools to compactly represent uncertainty distributions. Causal reasoning can be intuitively and formally described with graphs [2,3,6]. On these networks, it is possible to predict the effects of both observations and external actions on the system. Causal belief networks [3] are used to formalize the imperfect causal knowledge. They represent an alternative to causal Bayesian networks, that allow to formalize conditional beliefs in a flexible way. It is defined on two levels:

- Qualitative level: represented by a directed acyclic graph (DAG) named G where G = (V,E) in which the nodes V represent variables and edges E encode the cause-effect relations among variables. The set of parents of A is denoted by $Pa(A)$. The set of children of A is denoted by $Ch(A)$. A root is a node with no parents ($Pa(A) \neq \emptyset$). A leaf is a node with no children ($Ch(A) \neq \emptyset$). We will denote by R the set of roots and by L the set of leaves.
- Quantitative level: is the set of normalized *bba*s associated to each node in the graph. Conditional distributions can be defined for each variable A denoted on Θ_A in the context of its parents (either one or more than one node): $\sum_{sub_{ik} \subseteq \Theta_A} m^A(sub_{ik}|Pa(A)) = 1$

An intervention is an external action which changes some value(s) in the system and consequently will lead to different results than those found with observational data. These effects should be adequately predicted. While conditioning is used to compute the effect of observations, the "do" operator [6] is used to compute the impact of external action. Handling interventions and computing their effects on the system can be done by making changes on the structure of the belief causal network. The two equivalent methods developed were namely, belief graph mutilation method where all the edges directed to the target node will be deleted and belief graph augmentation method which consists of adding, for the target variable, a new parent variable denoted DO. Thus, the parents set of the variable A denoted PA is transformed to $Pa' = Pa \cup \{DO\}$. The DO node takes values in $do(a_i)$, $x \in \{\Theta_A \cup \{nothing\}\}$. do(nothing) represents the state of the system when no interventions are made. $do(a_i)$ means that the variable A is forced to take the certain value a_i.

4 DRC and GBT for Inference in Causal Belief Networks

4.1 Definitions

To reduce the cost of storage, the DRC and the GBT are used when the plausibility distributions are conditionally defined for singletons ($a_i \in a$) where $a \subseteq \Theta$. The DRC is used for backward propagation. Let m^Θ be the *bba* of the parent node A which is sent to its child C.

$$pl^\Omega(c) = \sum_{a \subseteq \Theta} m^\Theta(a)(1 - \prod_{a_i \in \Theta} (1 - pl^\Omega(c|a_i))) \quad (9)$$

The GBT can be used for forward propagation. Let m^Ω be the *bba* of the child node which is sent to its parent node A using the GBT.

$$pl^\Theta(a) = \sum_{c \subseteq \Omega} m^C(c)(1 - \prod_{a_i \in \Theta}(1 - pl^\Omega(c|a_i))) \tag{10}$$

Exemple 1. *Let $\Theta = \{a_1, a_2\}$ and $\Omega = \{c_1, c_2\}$. m^Θ and $m^\Omega(.|a_i)$ are the a priori mass distribution. pl^Ω is computed using the DRC (see Table 1). Thanks to the mobius transformation, we can convert the plausibility distribution pl^Ω to a mass distribution m^Ω. pl^Θ is computed using the GBT using m^Ω and $m^\Omega(.|a_i)$ (see Table 2).*

Table 1. DRC

	$\{a_1\}$	$\{a_2\}$	Θ	$pl^\Omega(c)$
$\emptyset$	0	0	0	**0**
$\{c_1\}$	0.5	0.2	0.6	**0.42**
$\{c_1\}$	0.5	0.8	0.9	**0.78**
Ω	1	1	1	**1**
m^Θ	0.2	0.4	0.4	

Table 2. GBT

	m^Ω	$\{a_1\}$	$\{a_2\}$	Θ
$\emptyset$	0	0	0	0
$\{c_1\}$	0.2	0.5	0.2	0.6
$\{c_1\}$	0	0.5	0.8	0.9
Ω	0.8	1	1	1
$pl^\Theta(a)$		**0.9**	**0.84**	**0.92**

4.2 Propagating Distributions for All Parents

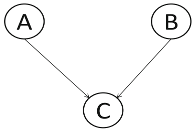
In this section, we explain that is possible to use the GBT and the DRC for propagation when the relations between nodes are not binary (i.e., conditional distributions are defined for all parents).

Given a set of variables $(A_1, A_2, \ldots, A_i)$ which are parent nodes of a variable C. To apply the GBT and the DRC, we will consider the m-tuple of the cartesian product of the parent nodes. Accordingly, the first component of the i-tuple belongs to A_1, the second A_2 and the i-th to A_i.

Each i-tuple will be considered as a singleton. To reduce the cost of storage and facilitate to experts to express their beliefs, the plausibility distributions of C will be defined and stored in the context of singletons of the cartesian product $(A_1, A_2, \ldots, A_i)$.

Example 2. *Let us consider the following directed causal belief network in Figure 1 where A and B are the parents of C. For the sake of simplicity, all the variables used in this example are binary. The DRC and the GBT can be applied for singletons ($a_i \in a$). Since the conditional mass distributions of C are defined for all parents A and B, conditional mass distributions will be defined for subsets. So, we have to use the cartesian product of the parents node A and B ($A \times B = \{a_1b_1, a_1b_2, a_2b_1, a_2b_2\}$). When applying the DRC and the GBT, the conditional plausibility distributions of C are saving according to singletons of the cartesian product $A \times B$.*

a_1	0.5
a_2	0.5
Θ_A	0



b_1	0.2
b_2	0.5
Θ_B	0.3

	a_1b_1	a_1b_2	a_2b_1	a_2b_2
c1	0.7	0.5	0.7	0.6
c2	0.2	0	0.3	0.4
Θ_C	0.1	0.5	0	0

Fig. 1. Belief network where distributions are defined for all parents

5 Inference in Singly Connected Causal Belief Networks

The impact of a new piece of information on the remaining variables can be found by first computing the joint distribution and then making marginalization by dropping the extra coordinates. This method is not suitable when the number of variables becomes substantial. To solve this problem, equivalent local computations have been proposed [5,7].

Existing algorithms only deal with the propagation of observational data in belief networks [1,12] where distributions are defined per single parent. To ensure propagation in the case where distributions are defined for all parents, we have to transform the initial network into a junction tree even if this technique is usually used to transform a multiply network into a tree structure. This transformation is expensive and the propagation algorithm depends on clusters' size. To tackle these problems, we propose a direct method of propagation in causal belief networks where distributions are defined for all parents. The proposed method consists of updating the belief mass of each node. If the node has more than one parent, we need to combine the distributions of parent nodes using the vacuous extension to the product space of variables representing the parent nodes. In our approach, the combined distribution is stored in a fictional node allowing message passing to its child node.

Causal propagation consists of finding the influence of an intervention or an observation on the remaining variables of the system. This is done through message passing between variables. When receiving a message each node X updates both local vectors; the vector $\pi(x_1, ..., x_n)$ concerning messages received by its parents and the vector $\lambda(x_1, ..., x_n)$ concerning messages received by its children. Each node sends and receives messages from each of its neighbors. The local message-passing between variables is based on two kinds of messages. The π-message is a message sent from a parent node to a child node and the λ-message is a message sent from a child node to a parent node.

In this section, we will first introduce the basic concepts of propagation in belief networks. Then, we explain how to compute the mass distribution of the fictional node. At the end of this section, we present algorithms for propagating observations and interventions in causal belief networks.

5.1 The Basic Concepts of Propagation in Belief Networks

Message passing is termed forward propagation or backward propagation depending on the direction in which the message is circulated. The causal belief inference algorithm are based on two rules the Disjunctive Rule of Combination (DRC) and the Generalized Bayesian Theorem (GBT). In fact, the algorithm consists of two phases: propagation down (backward) and propagation up (forward).

Let us consider the two nodes A and C where A is the parent of C. The message sent from A to C is a π-message computed using the DRC and the message sent from C to A is a λ-message computed using the GBT.

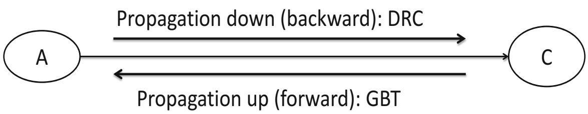


Fig. 2. Propagation process between A and C

5.2 Computation of the Mass Distribution of the Fictional Node

To perform the propagation in a simply connected network where nodes are originally defined in the context of all parents, we will combine the mass distributions of parent nodes. These distributions may correspond to the a priori distributions in the case of root nodes or posteriori distributions to the other nodes computed using the GBT and the DRC. The result of this combination will be stored in a table associated with a fictional node. The mass distribution of the fictional node is computed using the Dempster rule of combination after the extension of the mass distributions of different parents Pa(A) of the visited node A to a joint space using the vacuous extension. The mass distribution of the fictional node is denoted as $m_{fictional}$. It is computed as follows:

$$m_{fictional} = \oplus_{A \in Pa(A)} (m^{A \uparrow Pa(A)}) \tag{11}$$

Once we combined the distribution of the node parents, we can make the propagation up and down using the two operators DRC and GBT.

Example 2 (Continued). *Let us continue with the same network presented in Figure 1. Let m^A and m^B be the mass distributions of the two nodes A and B.*

To combine these nodes into a fictional node AB, we have to use the vacuous extension to extend A and B to a joint space $A \times B$ (see Table 3 and Table 4). The mass distribution of the fictional node is then computed using the Dempster rule of combination (see Figure 3).

Table 3. $m^{A \uparrow AB}$

$a_1 \times \Theta_B$	0.5
$a_2 \times \Theta_B$	0.5
$\Theta_A \times \Theta_B$	0

Table 4. $m^{B \uparrow AB}$

$b_1 \times \Theta_A$	0.2
$b_2 \times \Theta_A$	0.5
$\Theta_B \times \Theta_A$	0.3

$\{a_1,b_1\}$	0.1
$\{a_1,b_2\}$	0.25
$\{a_2,b_1\}$	0.1
$\{a_2,b_2\}$	0.25
$\{a_1\} \times \Theta_B$	0.15
$\{a_2\} \times \Theta_B$	0.15
$\Theta_A \times \{b_1\}$	0
$\Theta_A \times \{b_2\}$	0
$\Theta_A \times \Theta_B$	0

Fig. 3. A causal belief network with a fictional node

5.3 Propagation of Observations in Causal Belief Networks

We propose in this section a direct propagation algorithm in singly connected causal belief networks where conditional beliefs are defined for all parents. The causal direct propagation of observations consists of two steps: the propagation down and the propagation up. A post-order (in direction of leaves) and a pre-order (in direction of roots) will be defined to propagate information backward and forward respectively.

Algorithm. Propagation down

```
For each A ∈ post-order
  If A ∉ R
    Combine the masses of its parents using the vacuous extension.
    Store the combined distribution in a fictional node.
    Pass a message π from the fictional node to A using the DRC.
    Compute its mass distribution.
    Send a message to its child C.
    Marginalization: find the initial mass of the parent nodes.
  End if
End for
```

Algorithm. Propagation up

For each A $\in$ pre-order
  If A $\notin$ R
    Combine the masses of parents.
    Store the combined distribution in a fictional node.
    Send a message λ to the fictional node using the GBT.
    Compute its mass distribution.
    Marginalization: find the initial mass of the parent nodes.
  End if
End for

Algorithm. Direct propagation of observations
  Updating the mass distribution of the node concerned by the observation.
  Propagation down.
  Propagation up.

Each node A computes its mass distribution by combining the two values π and λ using this formula:

$$m \leftarrow \pi_A \oplus \lambda_A \tag{12}$$

Example 2 (Continued). *Let us continue with same example. Propagation consists of sending a message π from the mass distribution of the fictional node that is resulting from the combination of distributions of A and B to the node C. This latter computes its message π using the DRC (Equation 9). The new distribution of node C is as follows: c_1 = 0.73, c_2 = 0.17, Θ_C = 0.1*
Then, the node C sends a message λ to the fictional node AB which in turn computes the new value λ using the GBT (Equation 10). The results are subsets of the cartesian product of $A \times B$. The distribution of AB as follows:
$m^{AB}(\{(a_1, b_1)\}) = 0.0152$, $m^{AB}(\{(a_1, b_2)\}) = 0.0185$, $m^{AB}(\{(a_2, b_1)\}) = 0.026$, $m^{AB}(\{(a_2, b_2)\}) = 0.0404$, $m^{AB}(\{(a_1, b_1), (a_1, b_2)\}) = 0.0744$, $m^{AB}(\{(a_1, b_1), (a_2, b_1)\}) = 0.0065$, $m^{AB}(\{(a_1, b_1), (a_2, b_2)\}) = 0.01$, $m^{AB}(\{(a_1, b_2), (a_2, b_1)\}) = 0.0434$, $m^{AB}(\{(a_1, b_2), (a_2, b_2)\}) = 0.0279$, $m^{AB}(\{(a_2, b_1), (a_2, b_2)\}) = 0.0173$, $m^{AB}(\{(a_1, b_1), (a_1, b_2), (a_2, b_1)\}) = 0.1734$, $m^{AB}(\{(a_1, b_1), (a_2, b_1), (a_2, b_2)\}) = 0.0043$, $m^{AB}(\{(a_1, b_1); (a_1, b_2), (a_2, b_2)\}) = 0.1115$, $m^{AB}(\{(a_1, b_2), (a_2, b_1); (a_2, b_2)\}) = 0.065$, $m^{AB}(\{(a_1, b_1), (a_1, b_2), (a_2, b_1), (a_2, b_2)\}) = 0.3662$

After computing the distributions of the fictional node AB, it is possible to compute the mass distribution of A and B by applying the marginalization $m^{AB\downarrow A}$ where $m^{AB\downarrow A}(a_1) = 0.1081$, $m^{AB\downarrow A}(a_2) = 0.0837$ and $m^{AB\downarrow A}(\Theta_A) = 0.8082$ and $m^{AB\downarrow B}$ where $m^{AB\downarrow B}(b_1) = 0.0477$, $m^{AB\downarrow B}(b_2) = 0.0868$ and $m^{AB\downarrow B}(\Theta_B) = 0.8655$ since the propagation is ensured between a node and its neighbors.

5.4 Propagation of Interventions in Causal Belief Networks

Handling interventions can be done using graph augmentation and graph mutilation methods.

- **Propagation in the mutilated graph**

  Propagation in this graph consists of two steps: the mutilation step where the distribution of concerned by the intervention becomes a certain one (see Figure 4) and the propagation step using the direct causal propagation algorithm presented in Section 5.3.

Algorithm. Direct propagation using the mutilated based approach

Cutting all edges pointing to the node concerned by the intervention C.
C computes its new marginal which becomes a certain *bba*.
Propagation down.
Propagation up.

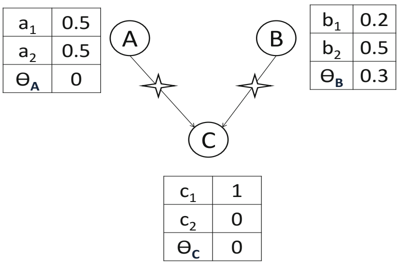


Fig. 4. A causal belief mutilated graph

- **Propagation in the augmented graph**

  Since adding the "DO" node, the conditional distribution of the node concerned by the intervention A given all parents must be updated. Hence, the graph augmentation method allows to represent the effect of observations when the DO node is taking the value nothing . When the DO node is taking the $do(a_i)$, we make a certain action which succeeds to put its target at a precise value by making it completely independent of its original causes. Thus, the distribution of A is a certain *bba*. Let Pa(A) be the parents of the A except the DO node, the conditional distribution of the A is defined as follows:

$$m(a_k|Pa(A), do(x)) = \begin{cases} 1 \text{ if } x = a_i \\ 0 \text{ if } x \neq a_i \\ m(a_k|Pa(A), do(x)) \quad x = nothing \end{cases} \tag{13}$$

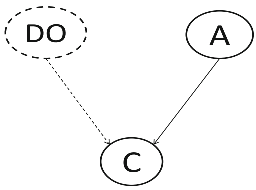
Propagation in this graph consists of two steps: the augmentation step where the conditional distribution of the node concerned by the intervention becomes a certain due to addition of the node DO and the propagation step using the using the direct causal propagation algorithm presented in Section 5.3.

Algorithm. Direct propagation using the augmented based approach

Add the node DO as a parent of the node concerned by the intervention C.
Updating the conditional mass distribution of C using Equation 13.
Propagation down.
Propagation up.

Example 3. *Let us consider the network presented in Figure 5 which illustrates a causal belief augmented graph on which an intervention $do(c_1)$ forces the variable C to take the specific value c_1. The conditional bba of C given its parents DO and A are defined using the Equation 13.*

$do(c_1)$	1
$do(c_2)$	0
do(nothing)	0



a_1	0.5
a_2	0.5
Θ_A	0

	$do(c_1)\ a_1$	$do(c_1)\ a_2$	$do(c_2)\ a_1$	$do(c_2)\ a_2$	$do(noth)a_1$	$do(noth)\ a_2$
c_1	1	1	0	0	0.7	0.5
c_2	0	0	1	1	0.3	0.5
Θ_C	0	0	0	0	0	0

Fig. 5. A causal belief augmented graph

In the case where the intervention $do(c_1)$ forces the variable C to take the specific value c_1, the mass distribution of the node C after the propagation process using the DRC is a certain bba. When the variable takes the value nothing, the bba is the same where there is no intervention. The parent nodes are extended to the joint space $DO \times A$ using the vacuous extension and then combined using the Dempster rule of combination. The results of the propagation are as follows: ***do***(c_1): c_1=1, c_2=0, Θ_C=0; ***do***(c_2): c_1=0, c_2=1, Θ_C=0; ***do***$(nothing)$: c_1=0.6, c_2=0.4, Θ_C=0

6 Conclusion

In this paper, we explained how we can still use the DRC and the GBT rules for propagation in causal belief networks even if the distributions are defined for all parents. We proposed a method acting directly on the network without having to go through the transformation into a junction tree. The proposed algorithms deal with the inference of observations and interventions in the augmented and mutilated graphs. As future work, we intend to treat inference in multiply connected causal belief networks. Inference in causal belief networks can be used in several applications like those allowing the intrusion detection and or ensuring system reliability.

References

1. Ben Yaghlane, B., Mellouli, K.: Inference in directed evidential networks based on the transferable belief model. International Journal Of Approximate Reasoning 48, 399–418 (2008)
2. Benferhat, S., Smaoui, S.: Possibilistic causal networks for handling interventions: A new propagation algorithm. In: AAAI Conference on Artificial Intelligence (AAAI), pp. 373–378. AAAI Press (2007)
3. Boukhris, I., Elouedi, Z., Benferhat, S.: Dealing with external actions in causal belief networks. International Journal Of Approximate Reasoning, 978–999 (2013)
4. Boussarsar, O., Boukhris, I., Elouedi, Z.: Representing interventional knowledge in causal belief networks: Uncertain conditional distributions per cause. In: Laurent, A., Strauss, O., Bouchon-Meunier, B., Yager, R.R. (eds.) IPMU 2014, Part III. CCIS, vol. 444, pp. 223–232. Springer, Heidelberg (2014)
5. Pearl, J.: Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann Pub., San Mateo (1988)
6. Pearl, J.: Causality: Models, Reasonning and Inference. Cambridge University Press (2000)
7. Shachter, R.D.: Probabilistic inference and influence diagrams. Operations Research 36, 589–604 (1988)
8. Shafer, G.: A Mathematical Theory of Evidence. Princeton Univ. Press, Princeton (1976)
9. Smets, P.: The combination of evidence in the transferable belief model. IEEE Pattern Analysis and Machine Intelligence 12, 447–458 (1990)
10. Smets, P.: Jeffrey's rule of conditioning generalized to belief functions. In: Uncertainty in Artificial Intelligence, pp. 500–505 (1993)
11. Xu, H., Smets, P.: Evidential reasoning with conditional belief functions. In: Uncertainty in Artificial Intelligence, pp. 598–606 (1994)
12. Xu, H., Smets, P.: Reasoning in evidential networks with conditional belief functions. International Journal of Approximate Reasoning 14, 155–185 (1996)