Causal Belief Networks: Handling Uncertain Interventions

Imen Boukhris[1,2], Salem Benferhat[2], and Zied Elouedi[1]

[1] LARODEC, Université de Tunis, ISG de Tunis, Tunisia
[2] CRIL, Université d'Artois, Faculté Jean Perrin, France
imen.boukhris@hotmail.com, benferhat@cril.univ-artois.fr, zied.elouedi@gmx.fr

Abstract. Eliciting the cause of an event will be easier if an agent can directly intervene on some variables by forcing them to take a specific value. The state of the target variable is therefore totally dependent of this external action and independent of its original causes. However in real world applications, performing such perfect interventions is not always feasible. In fact, an intervention can be uncertain in the sense that it may uncertainly occur. It can also have uncertain consequences which means that it may not succeed to put its target into one specific value. In this paper, we use the belief function theory to handle uncertain interventions that could have uncertain consequences. Augmented causal belief networks are used to model uncertain interventions.

1 Introduction

Despite its importance, causality is undefinable if a general and precise definition is sought (i.e., not restrained to particular cases) [22]. However, causal relations should be distinguished from mere statistical correlations. A paradigmatic assertion in causal relations is that the exterior manipulation (intervention) of a genuine cause will result in the variation of an effect. Therefore, interventions play a crucial role for an efficient causal analysis.

Bayesian networks [9,11,14] are successful graphical models representing a compact joint probability distribution. Causal Bayesian networks [14] go beyond Bayesian networks where arcs between variables follow the causal process. Probabilistic causal graphical models are effective when a very complete statistical knowledge description of the modeled system is available. If not, alternative causal networks will be more appropriate (possibilistic causal networks [3,4], causal belief networks [6]). On these networks, we can compute the simultaneous effect of observations and interventions. Interventions are distinguished from observations with the "do" operator [14]. An intervention forcing a variable A_i to be at a specific value a_{ij} is denoted by $do(a_{ij})$. This action deems that the original causes of the target variable are no more responsible of its state.

However, considering an intervention as a perfect external action is not realistic. Indeed, it may happen that due to an inattention, to ethical issues or to a lack of knowledge, the experimenter may not know the state of his action or its possible consequences. In fact, the occurrence of an intervention may be uncertain (e.g., injecting

L.C. van der Gaag (Ed.): ECSQARU 2013, LNAI 7958, pp. 61–72, 2013.

a drug whose expiration date has been exceeded). Moreover, an intervention may fail to set the target variable into one specific state (e.g., the use of a nicotine patches). In these cases, choosing random values will lead to the mis-estimation of the effects and accordingly to bad policies decisions.

Only few works in the probabilistic setting addressed the issue of intervention imperfection [10,12,20]. Besides in these works, interventions are defined differently from what is considered in the scope of this paper. In fact, they are considered as external actions certainly occurring represented with dummy variables that change the local probability distribution of the target variable.

The belief function theory is an uncertain framework that is especially appropriate to represent cases of partial and total ignorance. Therefore, it is an ideal tool to deal with these imperfect interventions. Despite its representation power, no work has been presented to handle uncertain interventions in the belief function framework.

This paper focuses on the modeling of uncertain interventions (i.e., uncertainly taking place) under the belief function framework. Graphically, to represent such interventions, augmented causal belief networks where conditional distributions are defined for any number of parents are used. In these networks, a conditional table is provided for the target variable given the intervention aside for the ones specified in the context of the initial causes. By this way, interactions with other causal factors are taken into consideration. Discounting technique is used to weaken the impact of the uncertain intervention on the distribution of the target variable. Moreover, a certain intervention may have uncertain consequences [7]. In this paper, we investigate the case of uncertain interventions that may have either certain or uncertain consequences.

The rest of the paper is organized as follows: in Section 2, we recall the basic concepts of the belief function theory and explain how causal knowledge can be represented on belief causal networks. The effect of uncertain interventions with certain consequences is handled in Section 3, whereas the case of uncertain interventions with uncertain consequences is treated in Section 4. Section 5 concludes the paper.

2 Belief Function Theory

2.1 Basics

We briefly recall the belief function theory. For more details see [15,19].

Let Θ be a finite set of mutually exhaustive and exclusive events referred to as the frame of discernment. The basic belief assignment (bba), denoted by m^Θ, is a mapping from 2^Θ to [0,1] such that:

$$\sum_{A \subseteq \Theta} m^\Theta(A) = 1 \tag{1}$$

When there is no ambiguity, m^Θ will be shortened m. The part of belief exactly committed to the event A of Θ is represented with the basic belief mass (bbm) denoted by $m(A)$. Subsets of Θ such that $m(A) > 0$ are called focal elements. When the emptyset is not a focal element, the bba is called normalized. 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.

Two *bbas* m_1 and m_2 induced by two distinct items of evidence can be aggregated using Dempster's rule of combination to give one resulting *bba* $m_1 \oplus m_2$.

$$m_1 \oplus m_2(A) = \begin{cases} K \cdot \sum\limits_{B \cap C = A} m_1(B) \cdot m_2(C), \;\; \forall B, C \subseteq \Theta & \text{if } A \neq \emptyset \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

where $K^{-1} = 1 - \sum\limits_{B \cap C = \emptyset} m_1(B) \cdot m_2(C)$ is the normalization factor.

The initial knowledge encoded with a mass value, $m(A)$, is revised using Dempster's rule of conditioning upon the arrival of a new certain piece of information B. All non vacuous events implying $\overline{B}$ will be transferred to the part of A compatible with the evidence namely, $A \cap B$ [17]. In the case, where $A \cap B = \emptyset$, several methods exist for transferring the remaining evidence [18]. $m(A|B)$ denotes the degree of belief of A in the context where B holds. It is defined as:

$$m(A|B) = \frac{\sum_{C, B \cap C = A} m(C)}{1 - \sum_{B \cap C = \emptyset} m(C)} \tag{3}$$

A basic belief assignment can be weakened (or discounted) before the combination to take into account the reliability of an expert by the discounting method defined as:

$$m^{\alpha}(A) = \begin{cases} (1-\alpha) \cdot m(A), \forall A \subset \Theta \\ \alpha + (1-\alpha) \cdot m(A), \;\; \text{if } A = \Theta \end{cases} \tag{4}$$

The discounting operation is controlled by a *discount rate* α taking values between 0 and 1. If α = 0, the source is fully reliable and beliefs remain unchanged. However, if α = 1, the *bba* is transformed into the vacuous *bba*, meaning that the information provided by the expert is completely discarded.

When a decision has to be made, beliefs held by the agent and represented by a *bba* could be transformed to a probability measure called $BetP$, using the pignistic transformation. It is defined as follows:

$$BetP(A) = \sum_{B \subseteq \Theta} \frac{|A \cap B|}{|B|} \frac{m(B)}{1 - m(\emptyset)}, \forall A \in \Theta \tag{5}$$

2.2 Causal Belief Networks

Belief networks [1,6,21] are simple and efficient tools to compactly represent uncertainty distributions. They have shown their efficiency in several applications (e.g., system analysis [16], threat assessment [2]). One main advantage of these networks is that they limit the use of a priori. They differ from Bayesian networks in the definition of conditional distributions and in the way to compute the global joint distribution. Causal belief networks [6,8] are seen as belief networks with some particular properties concerning the interpretation of arcs. They are defined on two levels as follows:

- qualitative level: a DAG $\mathcal{G} = (V, E)$ where arcs describe causal influence. Each variable A_i is associated with a finite set namely its frame of discernment Θ_{A_i} representing all its possible instances, i.e., $\{a_{ij,j=1,\ldots,|\Theta_{A_i}|}\}$. A variable A_j is called a parent of a variable A_i if there is an edge pointing from A_j towards A_i. The set of all parents of A_i is denoted by $U(A_i)$. Some of the parents of A_i are denoted by $PA(A_i)$ where a single parent is denoted by $PA_j(A_i)$. An instance from $U(A_i)$, $PA(A_i)$ or $PA_j(A_i)$ is denoted respectively by $u(A_i)$, $Pa(A_i)$ and $Pa_j(A_i)$.

- quantitative level: represented by the set of *bba*s associated to each node in the graph. For each root node A_i (i.e., $PA(A_i) = \emptyset$) having a frame of discernment Θ_{A_i}, an a priori m^{A_i} is defined on the powerset of $2^{\Theta_{A_i}}$, such that $\sum_{sub_{ik} \subseteq \Theta_{A_i}} m^{A_i}(sub_{ik}) = 1$. It is possible to model the total ignorance of the a priori by defining a vacuous *bba* on A_i (i.e., setting m(Θ_{A_i}) = 1). For the rest of the nodes, conditional distributions can be defined for each subset of each variable A_i in the context of its parents (either one or more than one parent node).

In causal belief networks, local conditional mass distributions are aggregated using the Dempster rule of combination. Since this rule is looking for intersections, each local distribution should be first extended to a joint frame. Thus, each conditional distribution will be deconditionalized (denoted by $\Uparrow$) and non-conditionalized distribution will be vacuously extended to a joint frame (denoted by $\uparrow$)[5].

$$m^{V=A_1,\ldots,A_n} = \oplus_{A_i \in V}(\oplus_{Pa_j(A_i)} m^{A_i}(a_i|Pa_j(A_i))^{\Uparrow A_i \times PA_j(A_i)})^{\uparrow V} \quad (6)$$

where the vacuous extension is computed as:

$$m^{A_i \uparrow A_i \times A_j}(a_i) = m^{A_i, A_j}(a_i \times \Theta_{A_j})$$

and a conditional distribution is deconditionalized as follows:

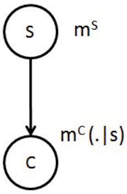
$$m^{A_i}(a_i|Pa_j(A_i))^{\Uparrow A_i \times PA_j(A_i)} = m^{A_i, A_j}(\{a_i \times Pa_j \cup \Theta_{A_i} \times \overline{Pa_j(A_i)}\})$$

On causal belief networks, it is possible to compute the effect of observations (seeing the natural behavior of the system) and interventions (intended external acting forcing a variable to take a specific value). If a manipulation of the event B leads to a change in A, then B is considered as a cause of A. While the effects of observations are computed with conditioning rules, those of interventions are handled by means of the so-called "do" operator [14]. An intervention in this case is considered as an external that totally control the state of its target variable. Such interventions make the original causes of the manipulated variable no more responsible of its state. All the other causes than the one of the intervention will be excluded. Graphically, interventions are described in two equivalent ways, namely graph mutilation and graph augmentation. The first way consists in modifying the causal graph by cutting off the links pointing into the target variable. The second equivalent way consists in adding, for the target variable, a new parent variable denoted DO.

3 Handling Uncertain Interventions with Certain Consequences

The occurrence of interventions recalled in the last section is assumed to be certain. However, it is not realistic to always consider interventions as fully certain external actions. An intervention having the variable A_i as target may uncertainly occur by forcing A_i to take *an unknown specific value* $a_{ij}(a_{ij} \in \Theta_{A_i})$ or it may *fail to take place.*

Example 1. *This example will be used in the rest of the paper to illustrate the main results. It concerns a description of knowledge regarding the causal link between the use of sugar and the sweetness of a coffee.*



$\{s_1\}$	0.1
$\{s_2\}$	0.7
Θ_S	0.2

	$\{s_1\}$	$\{s_2\}$
$\{c_1\}$	0.8	0.1
$\{c_2\}$	0.1	0.7
Θ_C	0.1	0.2

Fig. 1. Causal belief network

Fig. 1, depicts a causal belief network where S describes the presence of sugar in the cup of coffee, $\Theta_S = \{s_1, s_2\}$ where s_1 is yes and s_2 is no and C represents the sweetness of the coffee, $\Theta_C = \{c_1, c_2\}$ where c_1 is sweet and c_2 is bitter.

Let us assume that you have gone to a restaurant and ordered a coffee. A friend sees on the table a container with some white powder, without tasting it, he adds some of this powder into your cup of coffee because he knows that you like sweet coffee. Unfortunately, later he realizes that it may be either sugar or something else, and since you are in a restaurant it is most likely to be salt. If afterward, you taste the coffee and you find it sweet, you do not know if it is due to the action of your friend or to the way the coffee has been prepared. This latest alternative has no relation with the intervention of your friend. Thus, links relating the sweetness of the coffee with the initial use of sugar should not be deleted.

As in handling standard interventions, to represent uncertain interventions, we will alter the belief network by adding a new fictive node (DO) as a new parent of the variable A_i concerned by a manipulation, i.e., $PA(A_i) \leftarrow PA(A_i) \cup DO$. The DO node is taking value in $do(x)$, $x \in \{\Theta_{A_i} \cup \{\text{nothing}\}\}$. do(nothing) means that there are no actions on the variable A_i, it represents the state of the system when no interventions are made or totally fail to occur. do(a_{ij}) means that the variable A_i is forced to take the value a_{ij}. This way allows to represent the effect of interventions and also observations. The augmented graph is denoted by $\mathcal{G}_{aug}$. By taking advantage of the representation of causal belief networks to define conditional distributions [8], a conditional bba in the context of the fictive node DO will be "naturally" specified.

3.1 Interventions with an Unknown Specific Value

In the following, we propose a method to handle uncertain interventions that force the target variable to take an unknown specific value. To compute the distribution of the target variable, we need to address four different issues:

1- Deciding about the nature of the external action. We propose a general method where the nature of the intervention is undefined and we have to specify it. A *bba*, m^I, expressing the beliefs about the genuine nature of the external action expressed on a frame of discernment $\Theta_I = \{\theta_1, \ldots, \theta_n\}$ is defined. Note that the frame Θ_I may be different from the frame of the target variable. Deciding about the actual nature of the intervention will allow us to know which states will be affected by a change. The decision operation is made using the pignistic transformation.

Example 2 *(continued)***.** *Suppose that the beliefs about the nature of the substance in the container are flexibly expressed within the belief function formalism. They are defined on* $\Theta_I = \{sugar, salt, flour\}$ *such that* $m^I(\{sugar\}) = 0.2$, $m^I(\{salt\}) = 0.7$, $m^I(\{flour\}) = 0.01$ *and* $m^I(\{sugar,salt\}) = 0.09$. *The corresponding probabilistic knowledge of this bba is computed with the pignistic probability measure as follows:* $BetP^I(\{sugar\}) = 0.2 + 0.09 * 0.5 = 0.245$, $BetP^I(\{flour\}) = 0.01$, $BetP^I(\{salt\}) = 0.7 + 0.09 * 0.5 = 0.745$.

2- Defining the possible states of the intervention. The frame Θ_I is different from the frame of the target variable Θ_{A_i}. However, instances of Θ_I may affect the state of the target variable A_i by forcing it to take the value a_{ij}. Thus in the case of uncertain interventions, a matching between each θ_i and a state from Θ_{A_i} is defined as $match(\theta_i) = a_{ij}$. If θ_i has no impact on A_i, then we will say that $match(\theta_i) = nothing$. Note that more than one element of Θ_I may affect the same state a_{ij}.

Example 3 *(continued)***.** *The target variable has a frame of discernment* $\Theta_C = \{c_1 = sweet, c_2 = bitter\}$ *while the intervention is represented on* $\Theta_I = \{sugar, salt, flour\}$. *Table 1 presents the results of the matching between elements* θ_i *with instances of* C.

Table 1. Matching function: $match(\theta_i)$

θ_i	$match(\theta_i)$
sugar	c_1
salt	nothing
flour	c_2

Recall that the DO node represents the intervention. It has the same instances than its target to which the value $nothing$ is added. $do(a_{ij})$ means that the intervention attempts to set the target variable A_i into the state a_{ij}. This is achieved by performing the action θ_i. Therefore, executing θ_i amounts to $do(a_{ij})$. Accordingly, beliefs about the state of the variable DO reflecting the occurrence of the intervention will be defined from the knowledge about the decided nature of the intervention computed in the last step through BetPs. Since this latter reflects a probabilistic knowledge (i.e., computed for singletons), the *bba* of the DO node will be Bayesian and defined as:

$$m^{DO}(do(x)) = \begin{cases} \sum_{\theta_i, match(\theta_i)=a_{ij}} BetP^I(\theta_i) & \text{if } x = \{a_{ij}\} \\ \sum_{\theta_i, match(\theta_i)=nothing} BetP^I(\theta_i) & \text{if } x = \{nothing\} \end{cases} \quad (7)$$

Example 4 *(continued)***.** *According to the added substance, the coffee will be either sweet, bitter or remain as it was prepared. Therefore, forcing it to be at a specific state is not given for sure by adding the white powder. Hence, beliefs expressed about the actual occurrence of the intervention are computed using the BetP of each ingredient. In fact, the BetP takes into account all the focal elements that intersect with the substance of interest. The bba of the node DO is defined as:* $m^{DO}(\{do(c_1)\})$ = $BetP^I$*(sugar) = 0.245,* $m^{DO}(\{do(c_2)\})$ = $BetP^I$*(flour) = 0.01 and* $m^{DO}(\{do(nothing)\})$ = $BetP^I$*(salt) = 0.745.*

3- Defining Conditionals Given the DO Node. When occurring, an intervention $do(a_{ij})$ succeeds to force the variable A_i to take a certain value a_{ij}. Therefore, a conditional bba given an intervention is a certain bba focused on a_{ij} defined as:

$$m^{A_i}(sub_{ik}|do(a_{ij})) = \begin{cases} 1 & \text{if } sub_{ik} = \{a_{ij}\} \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

One can consider that $m^{A_i}(.|do(a_{ij}))$ is provided by an information source and this latest expects that it will be a certain bba. Since the occurrence of the intervention is uncertain, the bba defined by applying Equation 8 is not appropriate. Accordingly, this source is seen as not fully reliable. In fact, even if the intervention succeeds to put its target into one specific value, its occurrence remains uncertain. A Bayesian bba expressing the actual values concerning the occurrence of the intervention has been computed with BetP as explained in the last step. It will be used to evaluate the reliability of the source.

When considering the case of an intervention forcing the variable A_i to take the value a_{ij}, the occurrence of the intervention in the form of other states does not matter. What it was predicted by the source is an intervention certainly occurring at the state a_{ij}, $m^{DO}(do(a_{ij})) = 1$, whereas the actual belief about the occurrence of the intervention succeeding to put the variable A_i into the state a_{ij} is defined as $m^{DO}(do(a_{ij}))$ = $\alpha \in [0, 1]$. Since the degree of confidence in the reliability of a source can depend on the true value of the variable of interest, the difference between what is was predicted and the actual value is considered as its discounting factor defined as $1 - \alpha$. Consequently, the conditional distribution given the DO node is discounted by taking into account the reliability of each source, namely $\alpha_{do(a_{ij})}$. This information, will transform the conditional given the DO node from a certain bba into a weaker, less informative one. Hence, the new conditional bba of the target variable given the DO node becomes:

$$m^{A_i,\alpha_{do(a_{ij})}}(sub_{ik}|do(a_{ij})) = \begin{cases} 1-\alpha & \text{if } sub_{ik} = \{a_{ij}\} \\ \alpha & \text{if } sub_{ik} = \Theta_{A_i} \end{cases} \tag{9}$$

Proposition 1. *Standard interventions are a particular case of uncertain interventions when the source is fully reliable, i.e.,* $\alpha = 0$*.*

$$m^{A_i,\alpha_{do(a_{ij})}=0}(sub_{ik}|do(a_{ij})) = \begin{cases} 1 & \textit{if } sub_{ik} = \{a_{ij}\} \\ 0 & \textit{otherwise} \end{cases} \tag{10}$$

Example 5 *(continued)*. *Graphically, an extra node DO representing the intervention on the variable C is added as its new parent in the augmented graph. Each conditional distribution for the target variable C given an instance of the DO node is seen as provided by a distinct source of information. These sources affirm that performing an intervention will lead to a known change in the state of the manipulated variable. The conditional distributions as presented by the sources are presented in Table 2.*

Table 2. Certain bba: $m^C(.|do(x))$

	$\{do(c_1)\}$	$\{do(c_2)\}$	$\{do(nothing)\}$
$\{c_1\}$	1	0	0
$\{c_2\}$	0	1	0
Θ_C	0	0	1

Since the intervention achievement is uncertain, conditional local distributions presented in Table 2 are not appropriate. In fact, even when the intervention occurs with a degree of belief and succeeds to put its target into one specific value, one should take into consideration the cases where it fails to take place. Therefore, certain conditional local distributions will be discounted according to the reliability of each source. The degree of confidence in the reliability of a source is computed according the true value of the variable of interest, i.e., the DO bba. Hence, discount rates are denoted by $1-\alpha_{do(x)}$. *They are defined as 1-* $\alpha_{do(c_1)}$ *= 0.245, 1-* $\alpha_{do(c_2)}$ *= 0.01 and 1-* $\alpha_{do(nothing)}$ *= 0.745. The new discounted conditional bba is presented in Table 3.*

Table 3. Discounted bba: $m^{C,\alpha_{do(x)}}(.|do(x))$

	$\{do(c_1)\}$	$\{do(c_2)\}$	$\{do(nothing)\}$
$\{c_1\}$	1*0.245=0.245	0*0.01=0	0*0.745=0
$\{c_2\}$	0*0.245=0	1*0.01=0.01	0*0.745=0
Θ_C	0*0.245+0.755 =0.755	0*0.01+0.99=0.99	1*0.745+0.255=1

4- Defining Conditionals Given an Uncertain Intervention. The impact of the uncertain intervention on the target variable will not only depend from the intervention but also from the initial causes of the variable. To get the conditional bba given all the parent nodes, Dempster's rule of combination is used to aggregate the conditional distribution given the initial causes with the discounted conditional given the DO parent. We use $m^{A_i}(a_j|Pa(A_i))$ to represent the conditional mass function induced on the space Θ_{A_i} given $Pa(A_i) \subseteq \Theta_{PA(A_i)}$, and $m^{A_i,\alpha_{do(x)}}(a_k|do(x))$ to represent the discounted conditional mass function induced on the space Θ_{A_i} given the intervention $do(x)$. The bba of the target variable $m^{A_i}(a_i|Pa(A_i), do(x))$ is computed as follows:

$$m^{A_i}(a_i|Pa(A_i), do(x)) = \sum_{a_j \cap a_k = a_i} m^{A_i}(a_j|Pa(A_i)) \cdot m^{A_i,\alpha_{do(x)}}(a_k|do(x)) \quad (11)$$

Example 6 *(continued). The conditional bbas given the initial causes and that of the DO node can be aggregated to give the conditional bba* $m^C(.|s_i, do(x))$. *For instance,* $m^C(.|s_1, do(c_1))$ *is obtained by computing* $m^C(.|s_1) \oplus m^{C,\alpha_{do(c_1)}}(.|do(c_1))$. *Results are presented in Table 4.*

Unlike the case of standard interventions, $m^C(c_1|s_1, do(c_1)) \neq 1$. *However, the action of the friend has raised the beliefs about the sweetness of the coffee. A small increase from 0.8 to 0.845 is explained by the fact that it is more likely that the used ingredient is salt. In the same way,* $m^C(c_2|s_2, do(c_1))$ *has decreased from 0.7 to 0.638.*

Table 4. Conditional *bba*: $m^C(.|s_i, do(c_1))$

	$\{(s_1, do(c_1))\}$	$\{(s_2, do(c_1))\}$
$\{c_1\}$	0.8450	0.180
$\{c_2\}$	0.0775	0.638
Θ_C	0.0775	0.182

3.2 Interventions Not Occurring

The approach we proposed for handling interventions uncertainly happening remains valid to deal with the case of non-interventions. This is represented by setting the variable DO with certainty to the value $do(nothing)$.

In this paper, we consider that the situation of non-intervention encompasses:
- not acting on the target variable and observing the spontaneous behavior of the system,
- failing to act on the target variable and therefore the intervention will not occur.

Formally, in this case:

$$\forall \theta_i, match(\theta_i) = \{nothing\} \tag{12}$$

From Equations 7 and 12, the *bba* of the DO node is defined by:

$$m^{DO}(do(x)) = \begin{cases} 1 & \text{if } x = \{nothing\} \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

In this case, the state of the target variable will not depend on the intervention (i.e., from the DO node). The conditional *bba* given the DO node is not informative. It is represented with the vacuous *bba* defined as:

$$m^{A_i}(sub_{ik}|do(nothing)) = \begin{cases} 1 & \text{if } sub_{ik} = \Theta_{A_i} \\ 0 & \text{otherwise} \end{cases} \tag{14}$$

The "non-intervention" occurs certainly. Therefore, the source is fully reliable and the discounting factor is equal to zero. Hence, our approach well handles the particular case of standard interventions.

Proposition 2. *The beliefs provided about the non-occurrence of an intervention are accepted without any modification. They are defined like standard interventions.x*

$$m^{A_i,\alpha_{do(nothing)}}(.|do(nothing)) = m^{A_i}(.|do(nothing)) \tag{15}$$

The conditional bbas defined in the context of the DO node and of the initial causes are computed by combining each conditional defined per single parent as follows:

$$\begin{aligned} m^{A_i}(.|Pa(A_i), do(nothing)) &= m^{A_i}(.|do(nothing)) \oplus m^{A_i}(.|Pa(A_i)) \\ &= m^{A_i}(.|Pa(A_i)) \end{aligned} \quad (16)$$

Proposition 3. *An augmented causal belief graph where the DO node is set to the value $nothing$ encodes the same joint distribution than the initial causal belief network.*

$$m_{\mathcal{G}_{aug}}(.|do(nothing)) = m_{\mathcal{G}} \quad (17)$$

4 Handling Uncertain Intervention with Uncertain Consequences

In the last section, we dealt with interventions occurring in an uncertain way. When happening, even with a belief $m(\{do(a_{ij})\})$, they succeed to put the target variable into exactly one specific state. This situation is not always feasible. Therefore, our proposed approach in this section is to handle uncertain interventions with uncertain consequences, i.e., failing to put their target into a specific value.

4.1 Certain Interventions with Uncertain Consequences

In [7], we dealt with interventions that certainly take place but have uncertain consequences. To handle such cases, we proposed to specify a new bba on the target variable representing the consequences of the intervention. Let us denote by $\mathcal{F}_{A_i}$, the set of the focal elements representing the uncertain consequences of the intervention where a bbm β_j is allocated to each focal element. The conditional bba of the target variable given a certain intervention on the variable A_i attempting to force it to take the value a_{ij} is defined as follows:

$$m^{A_i}(sub_{ik}|do(a_{ij})) = \begin{cases} \beta_j & \text{if } sub_{ik} \in \mathcal{F}_{A_i}, \beta_j \in]0,1] \\ 0 & \text{otherwise} \end{cases} \quad (18)$$

Example 7 *(continued)***.** *Let us continue with the network of Fig. 1. Imagine here that your friend puts Lactose into your cup of coffee which is a disaccharide sugar. However it is known that it is poorly soluble. Therefore, even if the substance is a kind of sugar, adding it will obviously affect the sweetness of the coffee but without certainty. The conditional bba $m^C(.|do(c_1))$ defined upon this intervention is expressed as follows: $m^C(c_1|do(c_1)) = 0.8$, $m^C(c_2|do(c_1)) = 0.05$, $m^C(\Theta_C|do(c_1)) = 0.15$.*

4.2 Uncertain Interventions with Uncertain Consequences

In this paper, we also investigate the case of uncertain interventions with uncertain consequences. In fact, an intervention even taking place with a given degree of belief may have uncertain consequences. Remember that to deal with uncertain interventions succeeding to set their target into a specific value a_{ij}, the conditional bbas given instances of the DO node are discounted according to the actual occurrence of the intervention (see Equation 9). In the case of uncertain interventions with uncertain consequences, we take into consideration possible states that can take the target variable. Therefore, we define the resulting bba as a mixture of Equation 9 and 18 as follows:

$$m^{A_i}(sub_{ik}|do(a_{ij})) = \begin{cases} (1-\alpha)\cdot\beta_j & \text{if } sub_{ik} \in \mathcal{F}_{A_i} \\ \alpha + (1-\alpha)\cdot\beta_j & \text{if } sub_{ik} = \Theta_{A_i} \end{cases} \tag{19}$$

Proposition 4. *Uncertain interventions with a certain consequence are a particular case of uncertain ones with uncertain consequences when the parameter β_j is set to one.*

$$m^{A_i}(sub_{ik}|do(a_{ij})) = \begin{cases} 1-\alpha & \textit{if } sub_{ik} = \{a_{ij}\} \\ \alpha & \textit{if } sub_{ik} = \Theta_{A_i} \end{cases}$$

Example 8 *(continued)***.** *In context of a restaurant, it is more likely that what your friend has putted into your coffee is salt. We are focusing in the occurrence of the intervention as attempting to set its target into the value sweet, which means that the powder is sugar. However, some kinds of sugar (e.g., lactose, saccharine) are either not very soluble or may have a bitter or metallic unpleasant aftertaste. Adding them may lead to uncertain consequences. Note that the bbm that the added substance is sugar is represented with $m(\{do(c_1)\})$= 0.245. Hence, to represent this case the conditional bba given the DO node will be discounted. The resulting bba is presented in Table 5.*

Table 5. $m^{C,\alpha_{do(c_1)}}(.|do(c_1))$ upon an uncertain intervention with uncertain consequences

	$\{$do$(c_1)\}$
$\{c_1\}$	0.8*0.245 = 0.2
$\{c_2\}$	0.05*0.245 = 0.01
Θ_C	0.15*0.245+0.755 = 0.79

Note that as for uncertain interventions with certain consequences, the conditional distribution given the DO parent can be combined with the discounted conditional distribution given the initial causes using Dempster's rule of combination to obtain the conditional distribution given all the parent nodes.

5 Conclusion

This paper provided a causal graphical model to deal with interventions under the belief function framework. We argued that for several practical cases, interventions may be uncertain and should be consequently adequately modeled. Furthermore, we addressed the issue of uncertain interventions failing to be at one specific state so-called uncertain interventions with uncertain consequences.

We emphasized on that uncertain interventions have a natural encoding under the belief function framework and may be graphically modeled using causal belief networks. The effect of an uncertain intervention is computed on an altered structure, namely belief augmented graphs. In these networks, conditionals can be defined for any number of parents and are can be seen as provided by distinct sources of information.

As future works, we intend to explore the relationships between interventions and the belief changes using Jeffrey-Dempster's rule [13].

References

1. Ben Yaghlane, B., Mellouli, K.: Inference in directed evidential networks based on the transferable belief model. Int. J. Approx. Reasoning 48(2), 399–418 (2008)
2. Benavoli, A., Ristic, B., Farina, A., Oxenham, M., Chisci, L.: An application of evidential networks to threat assessment. IEEE Transactions on Aerospace and Electronic Systems 45, 620–639 (2009)
3. Benferhat, S.: Interventions and belief change in possibilistic graphical models. Artif. Intell. 174(2), 177–189 (2010)
4. Benferhat, S., Smaoui, S.: Inferring interventions in product-based possibilistic causal networks. Fuzzy Sets and Systems 169(1), 26–50 (2011)
5. Boukhris, I., Benferhat, S., Elouedi, Z.: Representing belief function knowledge with graphical models. In: Xiong, H., Lee, W.B. (eds.) KSEM 2011. LNCS, vol. 7091, pp. 233–245. Springer, Heidelberg (2011)
6. Boukhris, I., Elouedi, Z., Benferhat, S.: Modeling interventions using belief causal networks. In: FLAIRS 2011, pp. 602–607 (2011)
7. Boukhris, I., Elouedi, Z., Benferhat, S.: Dealing with interventions with uncertain consequences in belief causal networks. In: IPMU 2012, pp. 585–595 (2012)
8. Boukhris, I., Elouedi, Z., Benferhat, S.: On the modeling of causal belief networks. In: ICMSAO 2013 (to appear, 2013)
9. Darwiche, A.: Modeling and Reasoning with Bayesian Networks. Cambridge University Press (2009)
10. Eberhardt, F., Scheines, R.: Interventions and causal inference. Philos. Sci. 74, 981–995 (2007)
11. Jensen, F., Nielsen, T.: Bayesian Networks and Decision Graphs. Springer Publishing Company (2007)
12. Korb, K.B., Hope, L.R., Nicholson, A.E., Axnick, K.: Varieties of causal intervention. In: Zhang, C., Guesgen, H.W., Yeap, W.-K. (eds.) PRICAI 2004. LNCS (LNAI), vol. 3157, pp. 322–331. Springer, Heidelberg (2004)
13. Ma, J., Liu, W., Dubois, D., Prade, H.: Bridging jeffrey's rule, agm revision and dempster conditioning in the theory of evidence. International Journal on Artificial Intelligence Tools 20(4), 691–720 (2011)
14. Pearl, J.: Causality: Models, Reasonning and Inference. Cambridge University Press (2000)
15. Shafer, G.: A Mathematical Theory of Evidence. Princeton Univ. Press, Princeton (1976)
16. Simon, C., Weber, P., Evsukoff, A.: Bayesian networks inference algorithm to implement dempster shafer theory in reliability analysis. Reliability Engineering and System Safety 93, 950–963 (2008)
17. Smets, P.: The combination of evidence in the transferable belief model. IEEE Trans. Pattern Anal. Mach. Intell. 12(5), 447–458 (1990)
18. Smets, P.: About updating. In: UAI 1991, pp. 378–385 (1991)
19. Smets, P.: The transferable belief model for quantified belief representation, vol. 1, pp. 267–301. Kluwer Academic Publisher (1998)
20. Teng, C.M.: Applications of causal inference. In: ISAIM (2012)
21. Xu, H., Smets, P.: Evidential reasoning with conditional belief functions. In: UAI 1994, pp. 598–606 (1994)
22. Zadeh, L.: Causality is undefinable. Tech. rep., Univ. of California, Berkley (2001)