Possibilistic Causal Networks for Handling Interventions: A New Propagation Algorithm

Salem Benferhat and **Salma Smaoui**
CRIL,Universite d'Artois , Rue Jean Souvraz SP 18, 62300 Lens, France
{benferhat,smaoui}@cril.univ-artois.fr

Abstract

This paper contains two important contributions for the development of possibilistic causal networks. The first one concerns the representation of interventions in possibilistic networks. We provide the counterpart of the "DO" operator, recently introduced by Pearl, in possibility theory framework. We then show that interventions can equivalently be represented in different ways in possibilistic causal networks. The second main contribution is a new propagation algorithm for dealing with both observations and interventions. We show that our algorithm only needs a small extra cost for handling interventions and is more appropriate for handling sequences of observations and interventions.

Introduction

Bayesian probabilistic networks (Pearl 1988; Jensen 1996; Lauritzen & Spiegelhalter 1988) are powerful computational tools for identifying interactions among variables from observational data. Recently, Pearl (Pearl 2000) has proposed approaches based on probability theory using causal graphs to give formal semantics to the notion of interventions. From representational point of view, interventions are distinguished from observations using the concept of the "do" operator (Goldszmidt & Pearl 1992; Pearl 2000). From reasoning point of view, handling interventions consists of altering the graph by excluding all direct causes related to the variable of interest other than interventions by maintaining intact the rest of the graph. Effects of such interventions over the remaining variables are then computed by applying conditioning over the altered graph.
The "do" operator has been first proposed in (Goldszmidt & Pearl 1992) within ordinal conditional functions frameworks (Spohn 1988b; 1988a). Revisions and updates are unified through the conditioning on action to support causal reasoning. Spohn's ordinal conditional functions framework is a "qualitative" model for representing uncertainty, and has strong relationships with infinitesimal probabilities and with possibility theory (Dubois & Prade 1991).

This paper focuses on the developments of possibilistic causal networks in order to deal with both interventions and observations. It contains two main contributions:

- It provides theoretical foundations of possibilistic causal networks. Namely, it defines possibilistic causal networks and introduces the "do" operator in possibility theory. We show that the different equivalent representations of interventions in probabilistic networks are still valid in possibility theory framework.
- It provides a new propagation algorithm for dealing with both interventions and observations. We first show that a direct adaptation of propagation algorithms for probabilistic networks is not appropriate incrementally with new observations and interventions. Indeed, for multiply connected networks, when dealing with interventions, answering to queries is no longer immediate. It requires $O(|D|^{|u|\times|r|})$ computations where $|D|$ denotes the size of a variable domain, $|u|$ is the number of parents instances and $|r|$ is the number of different interventions. Now, if one imposes an efficient handling of queries, then a possible solution is to use augmented graphs (interventions are then encoded as observations on the augmented graph). This solution is not satisfactory, since for instance for multiply connected graphs, one cannot deal with sequences of observations and interventions unless the initialization step is repeated.

This paper takes advantage of properties of possibilistic networks and proposes a new algorithm where on one hand, junction tree is only constructed one time, and on the other hand queries are answered in a linear time.

The rest of the paper is organized as follows: next section gives a brief background on possibility theory and possibilistic networks. The possibilistic counterpart of the do operator is then proposed in the third section. Finally, we present our new possibilistic algorithm to deal with sequential non-simultaneous interventions. Last section concludes the paper.

Possibility Theory

This section only provides a brief background on possibility theory; for more details see (Dubois, Lang, & Prade 1994). Let $V = \{A_1, A_2, ..., A_n\}$ be a set of variables. D_{A_i}

denotes the finite domain associated with the variable A_i. a_i denotes any of the instances of A_i. X, Y, Z.. denote sets of variables. x is an element of the cartesian product $\times_{A_i \in X} D_{A_i}$ which is a subset of $\Omega = \times_{A_i \in V} D_{A_i}$ the universe of discourse. ω, an element of Ω, is called an *interpretation* or *event*. It is denoted by tuples $(a_1, ..., a_n)$, where a_i's are respectively instance of A_i's. $\omega[A_i]$ denotes the value that the variable A_i takes in the interpretation ω.

Possibility Measures and Possibility Distributions

A possibility distribution π is a mapping from Ω to the interval $[0, 1]$. The possibility degree $\pi(\omega)$ represents the compatibility of ω with available pieces of information. By convention, $\pi(\omega) = 1$ means that ω is totally possible, and $\pi(\omega) = 0$ means that ω is impossible. When $\pi(\omega) > \pi(\omega')$, ω is preferred to ω' for being the real state of the world. A possibility distribution π is said to be normalized if there exists at least one interpretation which is consistent with available pieces of information, namely: $\exists \omega \in \Omega, \ \ \pi(\omega) = 1$.
A possibility measure Π is a function that associates to each $\varphi \subseteq \Omega$ a weight in a unit interval $[0, 1]$. Π can be simply obtained from π as follows : $\Pi(\varphi) = \max\{\pi(\omega) : \omega \in \varphi\}$.
Conditioning (Hisdal 1978) consists of propositionally increasing elements consistent with x:

$$\pi(\omega \mid x) = \begin{cases} \frac{\pi(\omega)}{\Pi(x)} & if\ \omega[X] = x \\ 0 & otherwise. \end{cases} \quad (1)$$

Possibilistic Networks

Possibilistic networks (Fonck 1994; Borgelt, Gebhardt, & Kruse 1998), denoted by G, are directed acyclic graphs (DAG). Nodes correspond to variables and edges encode relationships between variables. A node A_j is said to be a parent of A_i if there exists an edge from the node A_j to the node A_i. Parents of A_i are denoted by U_{A_i} or U_i.
There are two kinds of possibilistic networks (depending on used conditioning). This paper focuses on possibilistic networks where conditioning is given by equation 1.
Uncertainty is represented at each node by local conditional possibility distributions. More precisely, for each variable A_i and for each u_i an element of the cartesian product of domains of variables which are parents of A_i, we provide $\Pi(a_i \mid u_i)$ for all $a_i \in D_{A_i}$, with $\max_{a_i \in D_{A_i}} \pi(a_i|u_i) = 1$.
Possibilistic networks are compact representations of possibility distributions. More precisely, the joint possibility distributions associated with possibilistic networks are computed using a so-called product-based possibilistic chain rule similar to the probabilistic one, namely :

$$\pi_{\Pi G}(a_1, ..., a_n) = \prod_{i=1,...,n} \Pi(a_i \mid u_i) \quad (2)$$

Example 1 *Figure 1 gives an example of possibilistic networks. The joint possibility distribution $\pi_G(AB) = \pi_G(A).\pi_G(B|A)$ associated with G is given in Table 1.*

Possibilistic Causal Networks

The ability of causal networks to predict the effects of interventions requires a stronger set of assumptions in

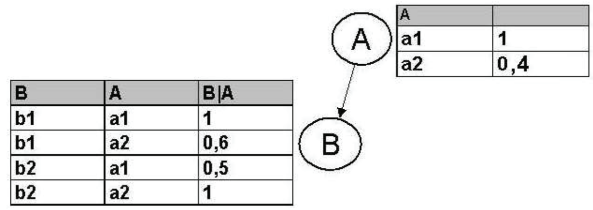

Figure 1: Example of a possibilistic causal network G

A	B	$\pi_G(AB)$
a_1	b_1	1
a_1	b_2	0.5
a_2	b_1	0.24
a_2	b_2	0.4

Table 1: The joint possibility distribution $\pi_G(AB)$

the construction of those networks. A possibilistic causal network is a possibilistic network where directed arcs of the graph are interpreted as representing causal relations between events. Arcs also follows the direction of causal process. Intuitively, the parent set U_i of A_i represents all the direct causes for A_i. Arrows indicate only that one variable is causally relevant to another, and say nothing about the way in which it is relevant.

The relation between causal and probabilistic information is a topic of several works. We argue that the same results also hold for other uncertainty theories since basically the main changes between probabilistic networks and probabilistic causal networks concerns the graphical structure. In this paper, we focus on possibilistic interpretations of causal relationships.
Interventions are handled as modalities over variables. A simple intervention, called "atomic", is one in which only a variable A_i is forced to take the value a_i. This intervention over A_i is denoted $do(A_i = a_i)$ or $do(a_i)$. $do(a_i)$ is viewed as an external intervention which alters just one mechanism (child-parent family) while leaving other mechanism intact.

The following subsection proposes a possibilistic model allowing to represent interventions using possibilistic causal networks.

Intervention as Negation of all Other Direct Causes

Pearl and Verma (Verma & Pearl 1990) interpreted the causal child-parent relationships in a DAG as a deterministic function $a_i = f_i(u_i, \gamma_i), i = 1, ..., n$ where u_i are the parents of variable A_i in a causal graph. $\gamma_i, 1 \leq i \leq n$ are independent disturbances and are instances of the unobservable variable set Γ .
The effect of an intervention, denoted $do(A_i = a_i)$ or $do(a_i)$, on Y (a subset of V) is induced from the model by deleting all equations associated with A_i and by substituting $A_i = a_i$ in the remaining equations.

In graphical models, interventions on a variable A_i also expresses that our beliefs (expressed in some uncertainty framework, here possibility theory) on parents U_i of A_i

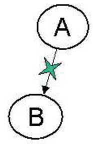

Figure 2: The mutilated graph G_m

should not be affected. This is represented by the deletion (also called mutilation) of the links between U_i and A_i. The rest of the graph remains intact. The resulting mutilated graph is denoted G_m such that $\pi(\omega|do(a_i)) = \pi_{G_m}(\omega|a_i)$, where π_{G_m} is the possibility distribution associated with the mutilated graph G_m.
The effect of the intervention $do(a_i)$ is to transform $\pi(\omega)$ into the possibility distribution $\pi_{a_i}(\omega) = \pi(\omega|do(a_i))$. We have,

$$\forall\omega, \pi_{a_i}(\omega) = \pi(\omega|do(a_i)) = \pi_{G_m}(\omega|a_i) \quad (3)$$

All other direct causes (parents) other than the action performed becomes independent of the variable of interest.

By observing the mutilated graph, the parents of the variable of interest becomes independent of the latter. Furthermore, the event that assigns the variable of interest A_i to the value a'_i after the performed intervention $do(a'_i)$ becomes certain. More formally, $\pi_{G_m}(a'_i) = 1$ and $\forall a_i \in D_{A_i}$: $a_i \neq a'_i, \pi_{G_m}(a_i) = 0$. Then, the effect of such intervention on the possibility distribution is given as follows, $\forall\omega$:

$$\pi(\omega|do(a'_i)) = \begin{cases} \prod_{j\neq i} \pi(a_j|U_j(\omega)) & if \;\; \omega[A_i] = a'_i \\ 0 & else \end{cases} \quad (4)$$

where $U_j(\omega) = u_j$ corresponds to the values that ω assigns to the parents of a_j.

Example 2 *Considering the possibilistic causal network G given in Example 1, the mutilated graph G_m obtained after intervention $do(B = b1) = do(b1)$ is given by the figure 2.*
The effect of the intervention $do(B = b_1)$ on the joint distribution $\pi(AB)$ associated with the graph in example 1 is given by the table 2

A	B	$\pi(AB\|do(b_1))$
a_1	b_1	1
a_1	b_2	0
a_2	b_1	0.4
a_2	b_2	0

Table 2: The joint possibility distribution $\pi(AB|do(b_1))$

This form of the equation (see 4) is interesting since it allows us to compute the effect of an intervention $do(a'_i)$ on a set of variable Y which is disjoint of $\{A_i \cup U_i\}$:

Proposition 1 *Let Y be a set of variable disjoint from $\{A_i \cup U_i\}$ where A_i is a variable in V altered by an intervention $do(a'_i)$ and U_i is the set of direct causes of A_i:*

$$\pi(y|do(a'_i)) = \max_{u_i} \pi(y|a'_i, u_i).\pi(u_i) \quad (5)$$

Proposition 1 is the counterpart of Pearl's proposition (Pearl 2000) in the probability theory framework. This result is not very surprising given the similarity between product-based possibilistic networks and a particular case of probabilistic networks where conditional probabilities are either close to 1 or close to 0.

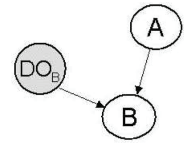

Figure 3: The augmented graph G_a

Adding Extra-Nodes

An alternative but equivalent approach (Pearl 1993) consists of considering the intervention as a new variable into the system. This subsection shows that this alternative approach is also valid in possibility theory. Intervention can be considered as a conditioning after altering the system. This alteration consists of adding a new link $DO_i \rightarrow A_i$ where DO_i represents the new intervention taking value in $\{\{do_{a_i} : \forall a_i \in D_{A_i}\}, do_{i-noact}\}$. The value "$do_{i-noact}$" (or "$do_{A_i-noact}$") means no intervention is performed on A_i. Values do_{a_i} mean that the variable A_i is forced to take the value a_i. The resulting augmented graph is denoted G_a. The parent set U_i is then augmented by the variable DO_i ($U'_i = U_i \cup DO_i$).
The new local possibility distribution at the level of the variable A_i after augmenting the graph is given by:

$$\pi(a'_i|u'_i) = \begin{cases} \pi(a'_i|u_i) & if \;\; DO_i = do_{i-noact} \\ 1 & if \;\; a'_i = a_i \\ 0 & if \;\; a'_i \neq a_i \end{cases} \quad (6)$$

As it is the case in probability theory, the two ways for handling interventions (mutilating or augmenting the graph) are also equivalent in possibility theory framework. Namely,

Proposition 2 *Let G be a possibilistic causal network and G_m (resp. G_a) be the mutilated (resp. augmented) graph obtained from G after performing intervention $do(a_i)$ on a variable A_i in G. Then, $\forall\omega$, $\forall a_i \in D_{A_i}$,*

$$\pi_{G_a}(\omega|DO_i = do_{a_i}) = \pi_{G_m}(\omega|A_i = a_i) = \pi(\omega|do(a_i))$$

This approach have the advantage to be applicable to any type of intervention. By adding a link $DO_i \rightarrow A_i$ to each node (on which intervention is possible) in the graph, we can construct a new possibility distribution containing information about richer types of interventions.

Example 3 *The augmented graph G_a obtained after intervention $do(b_1)$ from the graph G introduced in example 1 is given by the figure 3.*
The local possibility distribution $\pi(B|A, DO_B)$ at the level of B is given using 6. For instance, $\pi(b_1|a_2, do_{B-noact}) = \pi(b_1|a_2) = 0.6$ and $\pi(b_1|a_2, do_{b_2}) = 0$.

Propagation in Possibilistic Causal Networks

The above section has shown that the probabilistic handling of interventions can be easily adapted for possibility theory frameworks. This section provides a new propagation algorithm for reasoning from possibilistic causal networks which is better than a simple adaptation of the probabilistic one, since it allows an incremental handling of sequences of both observations and interventions.

Computing the effect of sequences of interventions and observations may be done either:

- by generalizing the explicit formula (5) to handle observations E (see 1) and a set of interventions F.
- or by adding a parent-node to each variable in F (augmenting the graph) and then applying standard possibilistic conditioning where interventions are interpreted as observations at the level of added nodes.

The generalization of (5) is given as a corollary of Proposition 1.

Corollary 1 *Let $E = e$ be an observation. Let F be the variables set affected by interventions and f be an element of $\times_{A_i \in F} D_{A_i}$. Let $U_F = \bigcup_{A_i \in F} U_i$, where U_i denotes the parents set of A_i. Let us denote by u_F an element of $\times_{A_i \in U_F} D_{A_i}$. The effect of the set of interventions denoted $do(f)$ and observations e on a remaining variable $A_j = a_j$ with $A_j \notin E \cup F \cup U_F$ is given by:*

$$\Pi(a_j|e, do(f)) = \max_{u_F} \pi(a_j|e, f, u_F).\pi(u_F|e) \quad (7)$$

Hence, to compute $\Pi(a_j|e, do(f))$, it is enough to compute for each u_F the expression $\pi(a_j|e, f, u_F).\pi(u_F|e)$, and then take the maximum of obtained results. The expression $\pi(a_j|e, f, u_F).\pi(u_F|e)$ can be obtained using any direct adaptation of probabilistic propagation algorithms since it corresponds to a simple conditioning. This is clearly not satisfactory, unless the number of interventions is low. Indeed, using such equation is inadequate when handling variables with important number of parents. Namely, this operation requires $O(|D|^{|u| \times |r|})$ where $|D|$ denotes the size of a variable domain, $|u|$ is the number of parents instances and $|r|$ is the number of interventions.

Another way to compute the effects of interventions consists of mutilating the initial graph. This approach is not adequate neither, especially for multiply connected graphs which require graphical transformation to junction trees in order to realize inference. In fact, let G be a multiply-connected possibilistic network. Suppose that an intervention $do(a'_i)$ is applied on the variable A_i. Then, the graph is mutilated and all arrows into A_i are deleted. A junction tree is formed after this operation. Suppose that another intervention $do(a'_j)$ is also applied on A_j such that $j \neq i$. By mutilating the initial graph, the junction tree resulting from G may change then it must be recomputed again.
The process of constructing junction trees is computationally expensive. This process must be repeated whenever changes are made to the network.
In the following, we present a new propagation algorithm using augmented graphs to deal with a sequence of interventions and observations.

A New Algorithm for Possibilistic Causal Networks

Our objective is to propose a new algorithm such that:

- queries are answered in a linear time.
- only a unique transformation from initial augmented graph into a junction tree is processed.
- observations and interventions may be handled incrementally (namely, the junction tree is incrementally updated without need of reinitialization).

The main point of our algorithm is the ability to express that by default there is no intervention, and then allowing to update possibility distributions if some interventions occur. The updating process should not lead to a degenerate case. Namely, we need to have an assignment of possibility distributions associated with new nodes added to the graph, i.e. we need to express on the "DO_i" nodes that there is no intervention by default. Unfortunately, in probabilistic networks, this is not easy to achieve (unless a reinitialization of the junction tree is repeated for each observation and intervention). In fact, let BN be a Bayesian network and let F be the variables set in G that may be directly affected by interventions. To each node A_i in F we assign a parent-node DO_i. The resulting graph is denoted BN_a.
One can expect that realizing inference without observing any evidence or applying any intervention on this augmented graph BN_a will induce an equivalent probability distribution (P_{BN_a}) to the initial one associated with BN (P_{BN}). Unfortunately, for any local probability distribution assigned to added nodes DO_i, including equiprobability, the a posteriori (i.e. resulting) distribution is different from the initial one except for local probability distributions that assign the highest degree (i.e. 1) to the value "$do_{i-noact}$" for each node DO_i and the degree 0 to all remaining values. More formally, $\forall \omega, \forall i : A_i \in F$,

$$P_{BN}(\omega) = P_{BN_a}(\omega) \text{ iff } P_{BN_a}(do_{i-noact}) = 1$$

Assigning such distributions (namely a degree 1 to $do_{i-noact}$) excludes any future interventions. Said differently, assigning the degree 0 to all other values different from "$do_{i-noact}$" means that it is impossible to have interventions on variables A_i.
In possibility theory, the situation is different, and this helps us to propose our new algorithm.

Local Possibility Distributions for DO variables

The following definition gives the local possibility distribution associated with a DO_i node, and which expresses that by default there is no intervention. This possibility distribution does not exclude future interventions.

Definition 1 *Let G be a causal network and G_a be the augmented graph obtained from G by adding a parent node DO_i for each variable A_i in F (variables set concerned by interventions). Let $A_i \in F$ be a node on which one may have an intervention. The a priori possibility distribution is defined on added node DO_i by:*

- $\pi_{G_a}(DO_i = do_{i-noact}) \leftarrow 1$,
- $\forall a_i \in D_{A_i}, \pi_{G_a}(DO_i = do_{a_i}) \leftarrow \epsilon$,
where ϵ is an infinitesimal (in fact, ϵ should be such that $\epsilon \leq \min_{\omega \in \Omega} \pi_G(\omega)$),

where π_G (resp. π_{G_a}) is the possibility distribution associated with the graph G (resp. G_a).

In fact, by assigning the possibility degree 1 to the "$do_{i-noact}$" for each added node DO_i, events $\{DO_i = do_{i-noact} : A_i \in F\}$ are accepted by default as an expected truth.
By assigning a degree ϵ to $\pi(DO_i = do_{a_i})$ for all instance $a_i \in D_{A_i} : A_i \in F$), events $\{DO_i = do_{a_i} : A_i \in F\}$ are considered to be the less normal and the less preferred in the graph, so that they do not bias our initial beliefs on the remaining (i.e. initial) variables.

The following proposition shows that if we are in a situation where there is neither interventions nor observations then the joint distributions associated with initial graph and augmented graph are the same. More precisely,

Proposition 3 *Let F be the set of manipulated variables. G_a denotes the augmented graph built from G by adding nodes DO_i as a parent to each $A_i \in F$ whose local distributions are given by Definition 1. The joint possibility π_{G_a} obtained from G_a over initial variables $V = \{A_1, ..., A_n\}$ is equivalent to the joint possibility distribution associated with the initial graph G. More formally,*

i) $\forall\omega\forall i : 1, ..., n, \pi_{G_a}(\omega) = \pi_G(\omega) = \pi_{G_a}(\omega|do_{i-noact})$,
ii) $\forall\omega\forall i : 1, ..., n, \pi_{G_a}(\omega|DO_i = do_{a_i}) = \pi_G(\omega|do(a_i))$.

Proof 1 *Let ω be an interpretation over the variables set $V = \{A_1, ..., A_i, ..., A_n\}$ and do_i be any instance of the variable DO_i. We have,*
i) $\pi_{G_a}(\omega) = \max_{do_i} \pi_{G_a}(\omega, do_i)$
$= \max_{do_i}(\pi_{G_a}(a_1|u_1)....\pi_{G_a}(a_i|u_i, do_i)....\pi_{G_a}(a_n|u_n).$
$\pi_{G_a}(do_i))$
$= [\prod_{a_j \neq a_i} \pi_{G_a}(a_j|u_j)].[\max(\pi_{G_a}(a_i|u_i, do_{i-noact}).$
$\pi_{G_a}(do_{i-noact}), \max_{do_{a'_i}}(\pi_{G_a}(a_i|u_i, do_{a'_i}).\pi_{G_a}(do_{a'_i})))]$
$= [\prod_{a_j \neq a_i} \pi(a_j|u_j)].[\max(\pi_{G_a}(a_i|u_i, do_{i-noact}),$
$\pi_{G_a}(do_{a_i}))]$
$= [\prod_{a_j \neq a_i} \pi(a_j|u_j)].[\max(\pi_{G_a}(a_i|u_i), \epsilon)]$
$= \prod_{a_i \in D_{A_i \in V}} \pi(a_i|u_i) = \pi_G(\omega)$
We also have,
$\pi_{G_a}(\omega|Do_i = do_{i-noact})$
$= \pi_{G_a}(a_1|u_1)...\pi_{G_a}(a_i|u_i, do_{i-noact})...\pi_{G_a}(a_n|u_n)$
$= \pi_{G_a}(a_1|u_1)...\pi_{G_a}(a_i|u_i)...\pi_{G_a}(a_n|u_n)$
$= \pi_{G_a}(\omega) = \pi_G(\omega)$
ii) When $DO_i \neq do_{i-noact}$, we obtain $\forall a_i : \omega[A_i] = a_i$,:
$\pi_{G_a}(\omega|Do_i = do_{a_i})$
$= \pi_{G_a}(a_1|u_1, do_{a_i})...\pi_{G_a}(a_i|u_i, do_{a_i})...\pi_{G_a}(a_n|u_n, do_{a_i})$
$= \pi_{G_a}(a_1|u_1)...\pi_{G_a}(a_i|u_i, do_{a_i})...\pi_{G_a}(a_n|u_n)$
Using equation 6 (definition of $\pi_{G_a}(a_i|u_i, do_{a_i})$), we obtain the same result as $\pi_G(\omega|do(a_i))$ (see 4).
When $\omega[A_i] \neq a_i$,
$\pi_{G_a}(\omega|Do_i = do_{a_i}) = \pi_G(\omega|do(a_i)) = 0.$
This result can be easily extended for handling several interventions.

Example 4 *Let us consider the possibilistic causal network G in the figure 1 and the augmented graph G_a in figure 3 after applying the intervention $do(B = b_1)$. Local possibility distribution at the level of the added node DO_B are given in table 3. Local possibility distribution at the level of B is computed using (6). The possibility distribution π_{G_a} related to G_a is given in table 4. The possibility distributions π_G (see table 1) and π_{G_a} over initial nodes A and B are equivalent. For instance, $\pi_G(a_2, b_1) = \pi_{G_a}(a_2, b_1) = 0.24$.*

DO_B	$\pi_{G_a}(DO_B)$
$do_{B-noact}$	1
do_{b_1}	0.001
do_{b_2}	0.001

Table 3: Local possibility distribution $\pi_{G_a}(DO_B)$

A	B	DO_B	π_{G_a}	A	B	DO_B	π_{G_a}
a_1	b_1	$noact$	1	a_2	b_1	$noact$	0.24
a_1	b_1	do_{b_1}	0.001	a_2	b_1	do_{b_1}	0.001
a_1	b_1	do_{b_2}	0	a_2	b_1	do_{b_2}	0
a_1	b_2	$noact$	0.5	a_2	b_2	$noact$	0.4
a_1	b_2	do_{b_2}	0.001	a_2	b_2	do_{b_2}	0.001

where $noact = do_{B-noact}$

Table 4: Joint distribution $\pi_{G_a}(A, B, DO_B)$

Augmented Junction Trees

A second important point is that the "DO_i" nodes can equivalently be added either before or after junction tree construction. From computational point of view it is better to first construct a junction tree associated with the initial possibilistic network. And, once the junction tree is constructed, we proceed to the addition of new nodes (namely DO_i). This is done when initializing the junction tree as follows:
- For each clique C_i in the junction tree: $\forall\omega, \pi_{C_i}(\omega) \leftarrow 1$ where π_{C_i} is the possibility distribution associated with C_i.
- For each node A_i in G, select a clique C_i containing $A_i \cup U_i$

- if $A_i \in F$ (a set of variables where intervention is possible), then
  - add the node DO_i to C_i
  - $\pi_{C_i} \leftarrow (\pi_{C_i}.\pi(A_i|U_i, DO_i).\pi(DO_i))$
- else $\pi_{C_i} \leftarrow (\pi_{C_i}.\pi(A_i|U_i))$

The following proposition shows that joint possibility distributions associated with the junction tree and the augmented graph G_a of G are equivalent.

Proposition 4 *Let G be a possibilistic causal network and F be a set of manipulated variables in G. do_F denote an element of $\times_{A_i \in F} D_{DO_i}$ and ω is an interpretation over $V = \{A_1, ..., A_n\}$. Let G_a be the augmented graph obtained from G adding to each node A_i in F a parent node DO_i. JT is the junction tree formed from G and initialized as indicated above. Then we have,*

$$\forall\omega\forall do_F, \pi_{JT}(\omega, do_F) = \pi_{G_a}(\omega, do_F)$$

where π_{JT} (resp. π_{G_a}) denotes the possibility distribution associated with JT (resp. G_a).

To summarize, our new propagation algorithm is described by the following steps:

- Construct a junction tree (JT) from the initial graph G.
- For each clique $C_i, \forall\omega, \pi_{C_i}(\omega) = 1$.

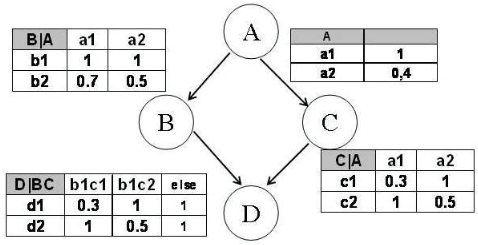

Figure 4: The multiply connected graph G

- For each node A_i in G, select a clique C_i containing $A_i \cup U_i$
  - if $A_i \in F$, then
    - add the node DO_i to C_i
    - $\pi_{C_i} \leftarrow (\pi_{C_i}.\pi(A_i|U_i DO_i).\pi(DO_i))$ (using Equation 6 and Definition 1),
  - else $\pi_{C_i} \leftarrow (\pi_{C_i}.\pi(A_i|U_i))$.
- Integrate evidence (observations and interventions).
- Propagate evidence until stability of the junction tree
- Answer queries

Example 5 *Let us consider the causal network in the figure 4. Let A and B be variables that may have interventions. The resulting augmented graph G_a is obtained by adding a parent node DO_A (resp. DO_B) to A (resp. B). Local possibility distribution assigned to added nodes DO_A and DO_B are given as follows: $\pi(do_{A-noact}) = \pi(do_{B-noact}) = 1$ and $\pi(do_{a_1}) = \pi(do_{a_2}) = \pi(do_{b_1}) = \pi(do_{b_2}) = 0.001$. The possibility degree associated to the event $\omega^+ = (a_1, b_1, c_1, d_1, do_{a_1}, do_{B-noact})$ is computed from G_a as follows:*

$$\begin{aligned}\pi_{G_a}(\omega^+) &= (\pi_{G_a}(a_1|do_{a_1}).\pi_{G_a}(b_1|a_1, do_{B-noact}).\pi_{G_a}(c_1|a_1).\\ &\quad \pi_{G_a}(d_1|b_1, c_1).\pi_{G_a}(do_{a_1}).\pi_{G_a}(do_{B-noact}))\\ &= 0.00009\end{aligned}$$

where local distributions at the level of A and B are computed using 6. Initializing the junction tree JT (formed from G) consists

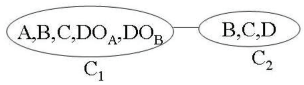

Figure 5: The junction tree JT after the initialization step

of initializing possibility distributions at the level of each clique:
$\pi_{C_1}(A, B, C, DO_A, DO_B) = (\pi(A|DO_A).\pi(DO_A).$
$\pi(B|ADO_B).\pi(DO_B).\pi(C|A)).$
$\pi_{C_2}(B, C, D) = \pi(D|BC).$
The junction tree obtained after initialization is given in figure 5. Computing the possibility degree of the event ω^+ from the initialized junction tree JT, we obtain:

$$\begin{aligned}\pi_{JT}(\omega^+) &= \pi_{C_1}(\omega^+).\pi_{C_2}(\omega^+)\\ &= \pi_{C_1}(a_1, b_1, c_1, do_{a_1}, do_{B-noact}).\pi_{C_2}(b_1, c_1, d_1)\\ &= 0.00009\end{aligned}$$

which is equal the possibility degree computed from the augmented graph G_a.

Conclusion

The first important contribution of this paper concerns theoretical foundations of possibilistic networks. We showed how interventions can be handled either by means of mutilated graphs or by means of augmented possibilistic causal networks or even by means of augmented junction trees. A new propagation algorithm through causal possibilistic networks dealing with both observations and interventions represents the second main contribution of this paper.

Acknowledgments

This work is supported by the project ANR Blanc MICRAC.

References

Borgelt, C.; Gebhardt, J.; and Kruse, R. 1998. Possibilistic graphical models. In *Proceedings of International School for the Synthesis of Expert Knowledge (ISSEK'98)*, 51–68.

Dubois, D., and Prade, H. 1991. Possibilistic logic, preferential models, non-monotonicity and related issues. In *IJCAI*, 419–425.

Dubois, D.; Lang, J.; and Prade, H. 1994. Possibilistic logic. In *Handbook on Logic in Artificial Intelligence and Logic Programming*, volume 3. OUP. 439–513.

Fonck, P. 1994. *Réseaux d'inférence pour le raisonnement possibiliste*. Ph.D. Dissertation, Université de Liège, Faculté des Sciences.

Goldszmidt, M., and Pearl, J. 1992. Rank-based systems: A simple approach to belief revision, belief update, and reasoning about evidence and actions. In *Proceeding of the Third International Conference (KR'92)*. Kaufmann. 661–672.

Hisdal, E. 1978. Conditional possibilities independence and non interaction. *Fuzzy Sets and Systems* 1:283–297.

Jensen, F. V. 1996. *Introduction to Bayesien networks*. University college, London: UCL Press.

Lauritzen, S. L., and Spiegelhalter, D. J. 1988. Local computations with probabilities on graphical structures and their application to expert systems. *Journal of the Royal Statistical Society* 50:157–224.

Pearl, J. 1988. *Probabilistic reasoning in intelligent systems: networks of plausible inference*. San Francisco (California): Morgan Kaufmman.

Pearl, J. 1993. Comment : Graphical models, causality and intervention. *Statistical Sciences* 8.

Pearl, J. 2000. *Causality: models, reasoning, and inference*. New York, NY, USA: Cambridge University Press.

Spohn, W. 1988a. A general non-probabilistic theory of inductive reasoning. In *Proceedings of the Fourth Conference on Uncertainty in Artificial Intelligence (UAI'88)*, 149–158.

Spohn, W. 1988b. Ordinal conditional functions: a dynamic theory of epistemic states causation in decision. In *Belief Changes and Statistics*. W. Harper and B. Skyrms. 105–134.

Verma, T., and Pearl, J. 1990. Equivalence and synthesis of causal models. In *UAI*, 255–270.