ECAI 2010
H. Coelho et al. (Eds.)
IOS Press, 2010

doi:10.3233/978-1-60750-606-5-943

Min-based causal possibilistic networks: Handling interventions and analyzing the possibilistic counterpart of Jeffrey's rule of conditioning

Salem Benferhat [1] and **Karim Tabia** [2]

Abstract. This paper deals with two important issues related to the handling of uncertain and causal information in a qualitative (or min-based) possibility theory framework. The first issue addresses encoding interventions using the possibilistic conditioning under uncertain inputs problem. More precisely, we analyze the min-based possibilistic counterpart of Jeffrey's rule of conditioning and point out that contrary to the probabilistic setting, this rule does not guarantee the existence of a solution satisfying the kinematics conditions. Then we show that this rule can naturally encode the concept of interventions in causal graphical models. Surprisingly enough, we show that when dealing with interventions the min-based counterpart of Jeffrey's rule provides a unique solution. The second issue deals with the efficient handling of sets of observations and interventions in min-based possibilistic networks, where we propose a solution based on a series of equivalent and efficient transformations on the initial causal graph.

1 INTRODUCTION

Possibility theory is among the main frameworks for representing and reasoning with uncertain information. Beliefs and background knowledge can be represented by means of possibility distributions and they can be revised and updated in the presence of new information. In probability theory, there are two main and similar methods for revising a probability distribution in case where the new information is uncertain. The first one is Jeffrey's rule [12] for revising probability distributions which is based on the probability kinematics principle. The second one is to use Pearl's method of virtual evidence [15] proposed in the context of probabilistic graphical models. The possibilistic counterpart of Jeffrey's rule was directly used in [10] without neither a real reference to probability kinematics nor an analysis of the existence/uniqueness of the solution. More recently, the possibilistic counterparts of Jeffrey's rule are investigated for possibilistic-based belief revision [3] and it is argued that this rule can successfully recover most of the belief revision kinds such as the natural belief revision [6], drastic belief revision [14], etc.

Possibilistic networks [1], in their two different forms (quantitative and qualitative), are graphical models based on possibility theory allowing to compactly represent the prior background knowledge and efficiently reason in the presence of new information. While the quantitative (or product-based) networks are very similar to Bayesian networks, the qualitative (or min-based) ones, which are the focus of this paper, have significant differences. Causal possibilistic models are updated with two types of information: observations and interventions which correspond to external actions forcing some variables to some values. Handling sets of observations and interventions is an important issue that can appear in many applications such as diagnosis and simulation where some pieces of information are directly observed (by testing some variables) while one must act on the system as an experimenter (by performing interventions) to obtain some other useful information. One way to handle sets of observations and interventions is to directly revise the joint possibility distribution encoded by the causal networks. This paper deals with these issues and addresses two important problems regarding the handling of uncertain and causal information in qualitative possibilistic networks:

- The first issue is related to the analysis of the existence and uniqueness of the solution obtained by the possibilistic counterpart of Jeffrey's rule of conditioning in the qualitative possibilistic framework. We point out that the kinematics conditions underlying Jeffrey's rule are too strong and does not always guarantee the existence of a solution. Then we show how interventions in possibilistic causal graphs can be naturally and equivalently handled using the possibilistic counterpart of Jeffrey's rule of conditioning. In particular, we show that the min-based possibilistic counterpart of Jeffrey's rule satisfies the requirements of handling interventions and leads to a unique revised possibility distribution when the uncertain inputs encode an intervention.
- The second issue deals with the efficient handling of sets of observations and interventions in qualitative possibilistic networks. We propose an efficient method for handling sequences of observations by directly performing equivalent transformations on the causal graph. This method takes inspiration in a method developed in [2] and guarantees the same results as handling observations by revising the underlying joint possibility distribution but without computing a new joint possibility distribution at each time an observation arrives. Finally, we point out that the proposed method for handling observations and interventions takes into account the order of arrival of interventions and observations.

2 BASIC BACKGROUNDS ON POSSIBILITY THEORY AND CAUSAL NETWORKS

Possibility theory is an alternative uncertainty theory [9] which uses a pair of dual measures to assess the knowledge relative to an event in hand. In the following, V=$\{A_1, A_2,..,A_n\}$ denotes a set of variables. D_{A_i}=$\{a_1, a_2,..,a_m\}$ denotes the domain of a variable A_i while Ω=$\times_{A_i \in V} D_{A_i}$ denotes the universe of discourse. An interpretation w=$(a_1, a_2, .., a_n)$ is an instance of Ω. ϕ, φ denote subsets

[1] CRIL UMR CNRS 8188 Artois University, France, email: benferhat@cril.fr
[2] LINA UMR CNRS 6241 Nantes University, France, email: Karim.Tabia@univ-nantes.fr

of Ω, called events. A possibility distribution π is a mapping from the universe of discourse Ω to the unit scale $[0,1]$ where a possibility degree $\pi(w_i)$ assesses to what extent it is consistent that the interpretation w_i can be the actual state of the world. By convention, $\pi(w_i)$=1 means that w_i is totally plausible and $\pi(w_i)$=0 denotes an impossible event. The relation $\pi(w_i)>\pi(w_j)$ means that w_i is more plausible than w_j. A possibility distribution π is normalized if $max_{w_i \in \Omega}(\pi(w_i))$=1. According to the interpretation underlying the possibilistic scale [0,1], there are two variants of possibility theory:

- **Qualitative (min-based) possibility theory**: In this case, the possibility distribution is a mapping from the universe of discourse Ω to an *ordinal* scale where only the *ordering* of values is important.
- **Quantitative (product-based) possibility theory**: In this case, the possibilistic scale [0,1] is numerical and possibility degrees can be manipulated by arithmetic operators.

In this paper, we only focus on the min-based setting which has significant differences with the probabilistic and product-based frameworks. Conditioning is a fundamental notion for updating a possibility distribution when a new evidence (consisting in a completely sure event) arrives. Note that there are several definitions of possibilistic conditioning [11][9] subject to the qualitative or quantitative interpretation underlying the possibilistic scale [0,1]. The qualitative possibilistic setting uses the well-known **min-based conditioning** proposed in [11][9] and defined as follows (we assume that $\Pi(\phi)\neq 0$):

$$\pi(w_i|\phi) = \begin{cases} 1 & \text{if } \pi(w_i)=\Pi(\phi) \text{ and } w_i \in \phi; \\ \pi(w_i) & \text{if } \pi(w_i)< \Pi(\phi) \text{ and } w_i \in \phi; \\ 0 & \text{otherwise.} \end{cases} \qquad (1)$$

In the min-based setting, conditioning a possibility distribution π with a sure input ϕ consists in making *impossible* all the interpretations that do not satisfy ϕ and changing only the possibility degree of the most plausible element of ϕ up to the value 1.

2.1 Causal possibilistic graphical models

Graphical models such as probabilistic networks [8][13] and possibilistic networks [4] are well-known and very efficient tools for representing and reasoning with uncertain, incomplete and complex information. Like Bayesian networks, possibilistic ones consist of two components (see the example of Figure 1):

- **A graphical component**: it consists in a directed acyclic graph (DAG) where the nodes denote the domain variables and arcs encode direct influence relationships existing between the variables.
- **A numerical component**: it is composed of a set of local possibility distributions measuring the influence endured by each variable A_i in the context of its parents U_i.

The normalization condition requires that every local possibility distribution should satisfy the following condition:

$$max_{a_i \in D_{A_i}}(\pi(a_i|u_i)) = 1. \qquad (2)$$

In the min-based possibilistic setting, the joint possibility distribution is factorized using the min-based chain rule:

$$\pi(a_1, a_2, .., a_n) = \min_{i=1}^{n} \pi(a_i|u_i). \qquad (3)$$

A *causal* possibilistic model refers to a possibilistic network where the graph only encodes causal relationships (each arc denotes a cause-effect relationship) instead of mere correlations. Hence, in a causal graph, the parent set U_i of a node A_i represents all the direct causes of A_i while A_i's children denote A_i's direct effects.

Example

In the rest of this paper, we will illustrate our results on a simplified example about a *toothaches* problem. Dental caries (or cavities) most often cause *toothaches* and *infections*. As shown in Figure 1, some *toothaches* are due to *gum* problems while caries result from inadequate *oral care* (and dietary habits). Note that this academic example is only concerned with patients consulting a doctor for toothaches. We define the following five variables:

- O (for *Oral care*) whose domain is D_O={*Good, Bad*}.
- C (for *Caries*) taking its values in D_C={*Yes, No*}.
- I (for *Infections*) taking its values in D_I={*Infected, NoInfected*}.
- T (for *Toothaches*) taking its values in D_T={*Aching, NotAching*}.
- G (for *Gum problems*) with D_G={*Healthy, NotHealthy*}.

The causal possibilistic network representing a doctor's beliefs on the toothaches problem is given in Figure 1 where the statement

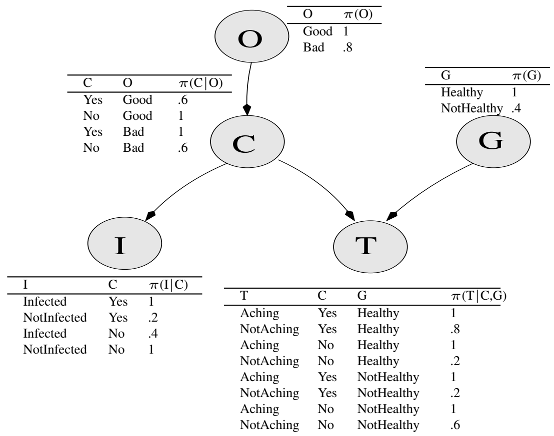

Figure 1. The causal possibilistic network of the toothaches problem

oral care is Good is completely plausible while the state Bad is exceptional. Similarly, $Healthy$ is the most common state for the gum problems variable G while the state $NotHealthy$ is very exceptional. Regarding the toothaches variable T, if the patient has a cavity (C=Yes) then the statement *tooth is aching* is completely plausible, etc.

Causal graphical models are expressive and compact representations which are updated with two forms of inputs: a set of evidences (observations) which are the results of *testing* some variables, and a set of interventions [16] which represent external actions that force some variables to have some specific values. One way to handle observations and interventions in causal graphical models is to view these tasks as belief change operations transforming the joint distribution encoded by the causal network into an a posteriori one constrained by the observation or the intervention.

3 JEFFREY'S RULE IN A QUALITATIVE POSSIBILISTIC FRAMEWORK

This section analyzes the conditioning under uncertain inputs and encoding interventions using Jeffrey's rule. In particular, we analyze an issue that has not been considered before regarding the uniqueness of conditioning rules with respect to the well-known kinematics rules in the context of min-based causal networks.

3.1 Qualitative possibilistic counterpart of Jeffrey's rule

Jeffrey's rule of conditioning [12] can be viewed as an extension of the standard probabilistic conditioning to the case where the evidence is uncertain [7]. The possibilistic counterpart of Jeffrey's rule is used for revising a $prior$ possibility distribution π into a new $updated$ or $posterior$ distribution π' given a set of $imposed\ constraints$ consisting in a set of mutually exclusive and exhaustive uncertain events λ_i. This method imposes two strong constraints corresponding respectively to the way the uncertain evidence is specified and the way the prior beliefs are revised:

1. **Specifying the uncertain evidence:** The uncertainty is of the form (λ_i, α_i) meaning that after the revision operation, the possibility degree of each event λ_i must be equal to α_i. The set $\{\lambda_1, .., \lambda_n\}$ induces a partition of Ω.
2. **Computing the revised possibility distribution:** The kinematics assumption states that even if there is a disagreement about the events $\lambda_1..\lambda_n$ in the old and new distributions, the conditional possibility degree of any event ϕ given any event λ_i should remain the same in the original and the revised distributions.

The two conditions underlying Jeffrey's rule of conditioning are formalized as follows:

Condition 1: $\forall\ \lambda_i{\subset}\Omega$, $\Pi'(\lambda_i)$=α_i
Condition 2: $\forall\ \lambda_i{\subset}\Omega$, $\forall\ \phi{\subseteq}\Omega$, $\Pi(\phi|\lambda_i)$=$\Pi'(\phi|\lambda_i)$.

In the following, we analyze the existence and uniqueness of the revised distribution π' in the min-based possibilistic framework.

3.2 Jeffrey's rule analysis in the min-based possibilistic setting

In the min-based setting, the revision according to Jeffrey's rule can be performed by the following formula (proposed in [10]):

$$\forall\phi \subseteq \Omega, \Pi'(\phi) = \Pi(\phi|(\lambda_i, \alpha_i)) = \max_{\lambda_i}(min(\Pi(\phi|\lambda_i), \alpha_i)). \quad (4)$$

Unlike the probabilistic and product-based possibilistic settings where there always exists a unique solution [5], in the min-based framework, there exist situations where Equation 4 does not guarantee the existence of a solution satisfying Condition 1 and Condition 2. As shown in the example of Table 1 and Table 2, this problem is due to the fact that in Equation 4, when imposing $\Pi'(\lambda_i)$=α_i lesser than $\Pi(\lambda_i)$, there might be several interpretations $w{\in}\lambda_i$ (namely those having $\pi(w){>}\alpha_i$) which will have their possibility degrees downgraded (and collapsed) to α_i in the revised distribution π' (namely, $\pi'(w)$=α_i). This results in losing the relative order of plausibility of interpretations w when moving from π to π'. One interesting option to be considered in the future is to replace the min operation in Equation 4 by an operation that encodes some lexicographic ordering.

Example(continued)

Assume that a doctor expresses his beliefs on dental cavities (variable C) and infections (variable I) using the joint possibility distribution π of Table1.

Assume now that following the results of recent and reliable surveys, the doctor wants to revise his beliefs in order to obtain $\Pi'(C{=}Yes)$=.2 and $\Pi'(C{=}No)$=1 (the doctor wants to decrease the possibility degree that an individual has a cavity). The revised possibility distribution π' computed using Equation 4 is given in Table 2. In Table 1, we observe that $\Pi(I{=}NotInfected|C{=}Yes)$=.6 while $\Pi'(I{=}NotInfected|C{=}Yes)$=1 in Table 2 violating Condition 2. In this example, Jeffrey's rule does not preserve the conditional possibility degree of having (or not) infections given that the individual has a cavity (or not).
Proposition 1 gives the exact conditions where the solution computed using Equation 4 does not satisfy Condition 1 and Condition 2.

C	I	$\pi(C,I)$	$\Pi(I\|C)$
Yes	$Infected$	.8	1
Yes	$NotInfected$	.6	.6
No	$Infected$	.4	.4
No	$NotInfected$	1	1

C	$\Pi(C)$
Yes	.8
No	1

Table 1. The possibility distribution π encoding the initial beliefs, the conditional distribution $\Pi(I|C)$ and the marginal distribution $\Pi(C)$

C	I	$\pi'(C,I)$	$\Pi'(I\|C)$
Yes	$Infected$	**.2**	1
Yes	$NotInfected$	**.2**	**1**
No	$Infected$	.4	.4
No	$NotInfected$	1	1

C	$\Pi'(C)$
Yes	**.2**
No	1

Table 2. The revised possibility distribution π', the underlying conditional distribution $\Pi'(I|C)$ and the marginal distribution $\Pi'(C)$

Proposition 1 Let π be the initial joint possibility distribution and (α_1, λ_1),..,(α_n, λ_n) be a set of uncertain inputs. Let $|\lambda_i, \alpha_i, \pi|$ denote the number of different plausibility levels of interpretations w satisfying λ_i such that $\alpha_i{\leq}\pi(w)$. Then

- If $\forall\lambda_i$, $|\lambda_i, \alpha_i, \pi|{<}2$, then there is a unique revised possibility distribution π' satisfying Condition 1 and Condition 2, computed using Equation 4.
- Otherwise (namely if $\exists\lambda_i$ such that $|\lambda_i, \alpha_i, \pi|{\geq}2$), using the min-based conditioning of Equation 1 there is no possibility distribution π' that satisfies Condition 1 and Condition 2.

Note that in case where Equation 4 does not guarantee the existence of a solution satisfying both Condition 1 and Condition 2, the revised possibility distribution π' still satisfies Condition 1.
In Jeffrey's rule view, the uncertain inputs are constraints that should be completely accepted (Condition 1) which may appear too strong in some applications. This however fully makes sense for handling interventions in graphical models as we will show in the following.

3.3 Encoding interventions with Jeffrey's rule

Handling interventions can be viewed as a belief change process that transforms a prior possibility distribution π encoding the initial beliefs into a new distribution π' constrained by the intervention. Assume that an intervention forces a variable A_i to take the value a_i^* (such an intervention will be henceforth denoted by $do(a_i^*)$ as in [16]). Assuming that the event A_i=a_i^* is somewhat plausible in π (namely, $\Pi(a_i^*){>}0$) and letting U_i be the set of direct parents of A_i in the causal graph G encoding π and D_{U_i} be the domain associated with U_i. Then, handling such an intervention in the new distribution π' should at least satisfy the two following constraints:

- **R1** the event A_i=a_i^* is a sure piece of information. Consequently, any interpretation $w{\in}\Omega$ where A_i is different from a_i^* is considered as completely impossible. Namely, $\forall w{\in}\Omega$, $\pi(w|do(a_i^*))$=0 if $w[A_i]{\neq}a_i^*$.

- **R2** the beliefs on the direct causes of A_i remain unchanged. Namely, $\forall A_j \in U_i$, $\forall a_j \in D_{A_j}$, $\Pi(a_j|do(a_i^*))=\Pi(a_j)$.

Then interventions can be naturally encoded using the possibilistic counterpart of Jeffreys rule by specifying the inputs as follows: When handling an intervention $do(a_i^*)$, the uncertainty bears on A_i's values. The revised distribution π' should then guarantee that $\Pi'(a_i^*)$=1 and $\Pi'(a_i)$=0 $\forall a_i \in D_{A_i}$ and $a_i \neq a_i^*$ (constraint R1) and the beliefs on U_i, the parents of A_i, must remain unchanged (constraint R2). By associating with each uncertain event $\lambda_i=\{a_i^*, u_i\}$ the plausibility degree $\alpha_i=\Pi(u_i)$ and for the remaining events $\{a_i\}$ where $a_i \neq a_i^*$ the degree 0, the obtained possibility distribution π' computed using Jeffrey's rule of Equation 4 satisfies the conditions R1 and R2 as formalized in the following proposition:

Proposition 2 Let π be a joint possibility distribution and $do(a_i^*)$ be an intervention forcing the variable A_i to take the value a_i^*. Let $\mu=\{((a_i^*, u_i), \Pi(u_i))$: $a_i^* \in D_{A_i}$ and $u_i \in D_{U_{A_i}}\} \cup \{(a_i, 0)$: $a_i \in D_{A_i}$ and $a_i \neq a_i^*\}$ be the set of exhaustive and mutually exclusive events. Then, the new possibility distribution π' (representing π revised with $do(a_i)$) obtained using Equation 4 satisfies the two constraints $R1$ and $R2$.

Proposition 2 states that encoding interventions using Jeffrey's rule fully satisfies the constraints R1 and R2 in the min-based framework.

3.4 Existence/uniqueness of the solution of Jeffrey's rule for encoding interventions

In Proposition 1, we pointed out that there exist situations where Jeffrey's rule of Equation 4 does not guarantee the existence of a solution satisfying both Condition 1 and Condition 2. Now, in the contexts of handling interventions, the revised distribution π' fully satisfies the constraints R1 and R2 and it is the unique one when the inputs represent interventions according to Proposition 2. This result is formalized in the following proposition:

Proposition 3 Let π be a joint possibility distribution and $do(a_i^*)$ be an intervention forcing the variable A_i to take the value a_i^*. Let π' be possibility distribution obtained using Equation 4 to encode the intervention $do(a_i^*)$ according to Proposition 2. Then π' satisfies the two constraints $R1$ and $R2$ and π' is unique.

The proof of Proposition 3 is straightforward since the revision using Jeffrey's rule of Equation 4 always satisfies Condition 1 (hence, the constraint R1 is satisfied) and one can easily check that the beliefs on parents of A_i are not altered in π' if the inputs are encoded according to Proposition 2. Moreover, since Proposition 3 entirely defines π', then π' is unique. Clearly, interventions provide a natural interpretation for Jeffrey's rule.

Let us now show that handling an intervention using the possibilistic counterpart of Jeffrey's rule of conditioning is equivalent to its handling by the graph mutilation method proposed in the context of causal graphical models [16][4]. Recall that in causal graphs, an intervention on a variable A_i, denoted $do(a_i^*)$, ensures that our beliefs on U_i (the set of parents of A_i) should not be affected. This can be achieved by deleting all the arcs from each variable composing U_i to A_i while maintaining the rest of the graph unmodified [16]. The obtained graph is called the $mutilated\ graph$ and denoted G_m such that $\pi_G(\omega|do(a_i^*))=\pi_{G_m}(\omega|a_i^*)$, where π_{G_m} (resp. π_G) is the possibility distribution associated with the mutilated graph G_m (resp. G). Now, in order to determine the effect of the intervention $do(a_i^*)$ on the rest of the initial graph G, one can apply the standard conditioning on the mutilated graph G_m after having observed the event A_i=a_i^*. Hence, the effect of this intervention on the joint possibility distribution is given by $\forall \omega$, $\pi_G(\omega|do(a_i^*))=\pi_{G_m}(\omega|a_i^*)$. Using Jeffrey's rule to encode the intervention $do(a_i^*)$ according to Proposition 2, the obtained a posteriori distribution π'_G is equivalent to $\pi_{G_m}(.|A_i=a_i^*)$ (namely, π_{G_m} conditioned with A_i=a_i^*) as formalized in the following proposition:

Proposition 4 Let G be a possibilistic network and π_G the joint distribution encoded by G. Let G_m be the mutilated graph obtained after handling an intervention and π_{G_m} the joint distribution encoded by G_m. Let π'_G be the possibility distribution obtained by revising π_G using the Equation 4 where the inputs encode the intervention $do(a_i^*)$ using Proposition 2. Then

$$\forall \omega \in \Omega,\ \pi_{G_m}(w|a_i^*)=\pi'_G(w).$$

Proposition 4 states that handling an intervention using Jeffrey's rule of conditioning induces the same possibility distribution as handling this intervention using the graph mutilation method. In the following, we propose an efficient method for handling sets of observations in qualitative causal possibilistic graphs.

4 HANDLING OBSERVATIONS IN QUALITATIVE POSSIBILISTIC NETWORKS

While handling interventions is straightforward and efficient in causal networks by mutilating the graph (or equivalently by augmenting the graph [16]), handling sets of observations using the standard conditioning is not efficient. Indeed, if π_G is the joint possibility distribution encoded by the possibilistic network G and A_i=a_i is an observation, then handling this observation requires computing the revised distribution $\pi_G(.|A_i=a_i)$. Clearly, revising the beliefs encoded by a possibilistic network G by revising the joint possibility distribution π_G encoded by G each time an observation is obtained is untractable. The solution proposed in [2] for handling observations in possibilistic graphs only deals with product-based networks (which are very similar to probabilistic networks). In the following, we show that in spite of the significant differences with respect to the product-based possibilistic setting especially regarding the conditioning and normalization operations, this solution can be adapted to efficiently handle sets of observations in qualitative possibilistic graphs. The main benefits of this approach is that unlike the standard conditioning which handles an observation by manipulating the whole joint distribution, the proposed solution performs this operation by altering only the necessary parts of the graph. The handling of an observation requires two steps: the first one allows to insert the new observation in the graph while the second step allows to renormalize the graph obtained after the first step.

Belief revision with observations[3] is traditionally done by a simple conditioning. Namely, if π_G is the joint possibility distribution encoded by the possibilistic network G and A_i=a_i^* is an observation, then the revised beliefs are π'_G=$\pi_G(.|A_i=a_i^*)$. Our goal is to propose a graphical counterpart for the the min-based conditioning operation. This graphical counterpart views conditioning as i) a *combination operation* followed by ii) a normalization operation. The combination operation combines the original possibility distribution with the one associated with the observation A_i=a_i^* while the normalization operation re-normalizes the possibility distribution obtained after the combination step in case where this latter becomes sub-normalized.

Let G be the causal possibilistic network encoding the initial beliefs

[3] In this paper, for each observation A_i=a_i and for each intervention $Do(A_i=a_i)$, we assume that A_i=a_i is somewhat possible (namely, $\Pi(A_i=a_i)>0$). See Section 4.2 for a brief discussion on belief change with impossible events.

and π_G be the possibility distribution encoded by G (π_G is obtained form G using the chain rule of Equation 3). In order to perform the combination operation, let us define the possibility distribution encoding the observation A_i=a_i as follows:

$$\forall \omega \in \Omega, \pi_{A_i=a_i^*}(\omega) = \begin{cases} 1 & if\ w[A_i] = a_i^* \\ 0 & otherwise \end{cases} \quad (5)$$

Clearly, in $\pi_{A_i=a_i}$ only the observed value a_i^* is totally possible while all the remaining ones are completely impossible which fully corresponds to the definition of an observation. Now, the combination of the initial beliefs with observation (namely, combining the possibility distribution π_G with $\pi_{A_i=a_i^*}$) can be defined as follows:

$$\forall \omega \in \Omega, \pi_{G2}(\omega) = \min(\pi_G(\omega), \pi_{A_i=a_i^*}(\omega)). \quad (6)$$

In Equation 6, the possibility distribution π_{G2} is obtained from π_G by considering as completely impossible every interpretation ω where the value of A_i is different from a_i^* (namely, $\forall\omega\in\Omega$, $\pi_{G2}(\omega)$=0 if $\omega[A_i]\neq a_i^*$), and preserving unchanged the possibility degrees of all interpretations ω where the value of A_i is a_i^*. After the combination step, the possibility distribution π_{G2} may be sub-normalized. Let us define the normalization operation as follows:

$$\pi_{G3}(\omega) = \begin{cases} 1 & if\ \pi_{G2}(\omega) = \Pi_{G2}(a_i^*) \\ \pi_{G2}(\omega) & otherwise \end{cases} \quad (7)$$

The normalization operation of Equation 7 upgrades the greatest possibility degree obtained after the combination operation such that the most plausible interpretation in π_{G2} becomes totally possible. Hence, using the combination (Equation 6) and normalization formulas (Equation 7), the min-based conditioning given by Equation (1) can be redefined as follows:

$$\forall \omega \in \Omega, \pi_G(\omega | A_i = a_i^*) = \pi_{G3}(\omega). \quad (8)$$

The following two subsections propose the graphical counterparts for the combination and normalization operations. For lack of space, we restrict ourself to possibilistic graphs where DAG's are trees (where a node has at most one parent).

4.1 Inserting the observation in the graph

Handling an observation A_i=a_i^* in a causal possibilistic graph $G1$ requires that in the obtained network $G2$ the value a_i^* is associated with a possibility degree $\Pi_{G2}(a_i^*)$=1 and $\forall a_i\in D_{A_i}$ such that $a_i\neq a_i^*$, $\Pi_{G2}(a_i)$=0. Moreover, an observation regarding the variable A_i alters the beliefs on A_i's parent denoted U_i. Let us use $G2$ to denote the result of inserting the new observation A_i=a_i^* in the network $G1$. The network $G2$ is specified as follows:

Definition 1 The possibilistic network $G2$ resulting from inserting the observation A_i=a_i^* in the network $G1$ is defined as follows:

- the structure of $G2$ is obtained from the DAG of $G1$ by deleting the arc from the parent variable U_i to A_i.
- the local possibility distribution of any variable A_k in $G2$ different from A_i and its parent U_i is identical to A_k's local distribution in $G1$.
- the new local possibility distributions of variable A_i and its parent denoted U_i, are defined as follows:
  - $\forall a_i\in D_{A_i}$,
  $$\pi_{G2}(a_i) = \begin{cases} 1 & if\ a_i = a_i^* \\ 0 & otherwise \end{cases}$$
  - Let A_j be the parent of A_i (if any) and U_j the parent of A_j (if any), then $\forall a_j\in D_{A_j}$ $\forall u_j\in D_{U_j}$,
  $$\pi_{G2}(a_j \mid u_j) = \min(\pi_{G1}(a_j \mid u_j), \pi_{G1}(a_i^* \mid a_j))$$

It is straightforward that in the new local possibility distribution relative to the variable A_i only the instance a_i^* is completely plausible and all the other instances are completely impossible. Hence, since the value of the variable A_i is now fully determined, there is no need to maintain the arc from the parent node U_i to A_i. Unlike interventions, an observation regarding a variable A_i alters the beliefs on its parent U_i. Accordingly, the distribution of U_i are altered as formulated in the following proposition:

Proposition 5 Let $G2$ be the possibilistic network obtained from $G1$ using Definition 1.
Then $\forall\omega\in\Omega$, $\pi_{G2}(\omega)$=$\min(\pi_{G1}(\omega), \pi_{A_i=a_i^*}(\omega))$.

Let us illustrate the transformation of Definition 1 on our example:

Example (continued)

We continue with the example of Figure 1 but restricted to a tree by discarding the nodes C, I and O. Figure 2 gives the initial network $G1$ and $G2$ obtained after combining the network $G1$ with the observation T=$Aching$.
As a result of inserting the observation A_i=a_i^*, the new local distributions of variable U_i (the parent of the observed variable) may become sub-normalized, the following step deals with this problem.

4.2 Re-normalizing the graph

After inserting an observation, the possibility distribution of the parent node A_j of the observed one A_i may be sub-normalized. Namely, it may exists an instance u_j of the parent variable of A_j denoted U_j such that $\max_{a_j\in D_{A_j}}(\pi_{G2}(a_j|u_j))$=$\beta$ with β<1. The re-normalization step allows computing a new possibilistic network $G3$ such that $G2$ and $G3$ encode the same joint distribution (namely $\forall\omega\in\Omega$, $\pi_{G2}(\omega)$=$\pi_{G3}(\omega)$) while all the local distributions in $G3$ are normalized. $G3$ is obtained as follows:

Definition 2 Let $G2$ be the network obtained using Definition 1 where the observation A_i=a_i^* is inserted in the initial causal graph $G1$. Let A_j be the parent of A_i whose possibility distribution is sub-normalized after Step 1. Let U_j be the parent of A_j. Let also u_j^* be the instance of U_j such that $\max_{a_j\in D_{A_j}}(\pi_{G2}(a_j|u_j^*))$=$\beta$ with 0<β<1 and let a_j^*=$argmax_{a_j\in D_{A_j}}(\pi_{G2}(a_j|u_j^*))$. The network $G3$ is such that it has exactly the same DAG as $G2$ and

- $\forall A_k$, $A_k\neq A_j$ and $A_k\neq A_i$ and $A_j\neq U_j$, $\pi_{G3}(a_k \mid u_k)$=$\pi_{G2}(a_k \mid u_k)$,
- $\forall a_j\in D_{A_j}, \forall u_j\in D_{U_j}$,
$$\pi_{G3}(a_j|u_j) = \begin{cases} 1 & if\ u_j = u_j^* and\ a_j = a_j^* \\ \pi_{G2}(a_j|u_j) & otherwise \end{cases}$$
- Let U_k be the parent node of U_j (if any) in $G2$. $\forall u_j\in D_{U_j}$, $\forall u_k\in D_{U_k}$,
$$\pi_{G3}(u_j|u_k) = \begin{cases} \min(\pi_{G2}(u_j|u_k), \beta) & if\ u_j = u_j^* \\ \pi_{G2}(u_j|u_k) & otherwise \end{cases}$$

The transformation of Definition 2 ensures that the networks $G2$ and $G3$ encode the same joint possibility distribution (namely, $\forall\omega\in\Omega$, $\pi_{G2}(\omega)$=$\pi_{G3}(\omega)$). It alters only the local distributions of A_j and its patent U_j. More precisely, it first normalizes the local possibility distribution of A_j then performs the inverse operation of normalization on the possibility distribution of U_j in order to ensure that the joint possibility distribution encoded by $G2$ and $G3$ remain the same. The case where A_j is a root node is a special case of Definition 2. Indeed, if A_j is a root node then the re-normalization is achieved by assigning a possibility degree of 1 to the most plausible instance of A_j in the network $G2$. As a side effect of re-normalizing A_j's distribution, its parent's distribution may become in turn sub-normalized. Then the re-normalization process should be repeated until reaching the root node. Let us illustrate the transformations of Definitions 1 and 2 on the example of Figure 2.

Example (continued)

Here, the initial network $G1$ is a part of the network of Figure 1 limited to variables T, C and O only for simplicity's sake. Figure 2 shows that the local distribution relative to the non root node C of network $G2$ (obtained after inserting the observation $T{=}Aching$ in the network $G1$ using Definition 1) is sub-normalized. The normalization of this distribution according to Definition 2 gives the network $G3$ of Figure 2. Now the normalization of C renders O sub-normalized. This latter is normalized also using Definition 2 giving the network $G3'$.

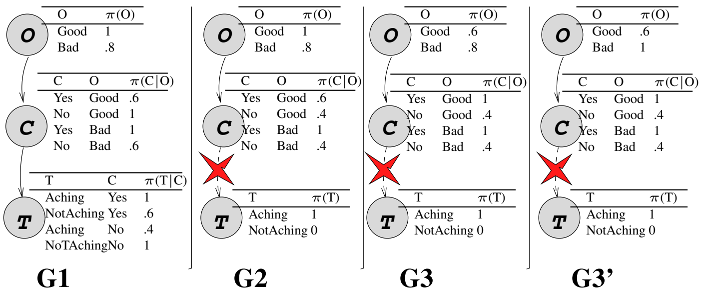

Figure 2. The initial network $G1$ and $G2$ (resp. $G3$ and $G3'$) obtained after applying Definition 1 (resp. Definition 2)

One can easily check that the joint distribution encoded by network $G3'$ is exactly the same as the one encoded by $G2$ while $G3'$ is completely normalized. Regarding the computational complexity of our two-steps procedure for handling observations in min-based causal possibilistic graphs, we argue that it is in linear time with respect to the initial network's parameters (number of local and conditional possibility degrees in the network) for tree structures.

In the previous section, we assumed that the events should be somewhat possible in the initial beliefs. This fully makes sense in case of observations since we are in presence of static world. However, the situation is different when one deals with interventions. For instance, assume that we have a causal network where there is a variable C (color of the fence) which is set with a full certainty to the value $blue$. Now, assume that we have an intervention that forces the color of the fence to the value red. Clearly, there is absolutely no contradiction in the presence of this intervention. However, this can be hardly managed if one directly achieves this change on possibility distributions using the possibilistic conditioning. Now, in the causal possibilistic network, this can be easily achieved by simply changing the prior possibility distribution of the color node C. Again, the structure of the causal network is crucial for handling belief change by impossible events in presence of interventions. Finally, it is worth pointing out that unlike handling sets of only observations (resp. interventions) where the order of observations (resp. interventions) does not matter, the situation is different when handling sequences involving both observations and interventions. This issue is not surprising since in the context of belief revision, it is well-known that Jeffrey's rule is not commutative. This problem also occurs in the min-based possibilistic framework where given a qualitative possibilistic network encoding the initial beliefs, there might exist situations where the revised beliefs after having an observation followed by an intervention will not give the same results as if we have first the intervention then the observation. Note lastly that handling observations using the method proposed in Section 4 and the graph mutilation (or augmentation) method [16] for handling interventions takes into account the order of arrival of observations and interventions.

5 CONCLUSIONS

This paper dealt with two important issues regarding the handling of uncertain and causal information in qualitative causal possibilistic networks. In particular, we addressed an issue that has not been investigated before regarding the existence and uniqueness of the well-known Jeffrey's rule in a qualitative possibilistic framework. We provided the exact conditions where a solution exists. We also showed that the strong kinematics constraints underlying this rule provide a natural way for encoding interventions and we argue that when used for this purpose, Jeffrey's rule provides always a unique solution. The second important issue addressed in this paper is related to the handling of sets of observations and interventions. We proposed an efficient method for handling sets of observations which first integrates the new observation in the causal graph then proceeds to the re-normalization of sub-normalized distributions if any. Future directions will consist in comparing Jeffrey's rule with Pearl's method of virtual evidence [15] in a qualitative possibilistic framework.

ACKNOWLEDGEMENTS

This work is supported by the (ANR) SETIN 2006 PLACID project (`http://placid.insa-rouen.fr/`)

REFERENCES

[1] N. Ben-Amor, S. Benferhat, and K. Mellouli, 'Anytime propagation algorithm for min-based possibilistic graphs', *Soft Comput.*, **8**(2), 150–161, (2003).

[2] S. Benferhat, 'Interventions and belief change in possibilistic graphical models', *Artificial Intelligence*, **174**(2), 177 – 189, (2010).

[3] S. Benferhat, H. Prade D. Dubois, and M.-A. Williams, 'A framework for iterated belief revision using possibilistic counterparts to jeffrey's rule', *Fundam. Inform.*, **99**(2), 147–168, (2010).

[4] S. Benferhat and S. Smaoui, 'Possibilistic causal networks for handling interventions: A new propagation algorithm', in *The 22nd AAAI Conference on Artificial Intelligence*, pp. 373–378, (2007).

[5] S. Benferhat, K. Tabia, and K.Sedki, 'On analysis of unicity of jeffrey's rule of conditioning in a possibilistic framework', in *11 th International Symposium on Artificial Intelligence and Mathematics*, Florida, USA, (2010).

[6] C. Boutilier, 'Revision sequences and nested conditionals', in *Thirteenth International Joint Conference on Artificial Intelligence*, pp. 519–531, (1993).

[7] H. Chan and A. Darwiche, 'On the revision of probabilistic beliefs using uncertain evidence', *Artif. Intell.*, **163**(1), 67–90, (2005).

[8] A. Darwiche, *Modeling and Reasoning With Bayesian Networks*, Cambridge University Press ELT, New York, NY, USA, April 2009.

[9] D. Dubois and H. Prade, *Possibility Theory: An Approach to Computerized Processing of Uncertainty*, Plenium Press, 1988.

[10] D. Dubois and H. Prade, 'A synthetic view of belief revision with uncertain inputs in the framework of possibility theory', *Int. J. of Approximate Reasoning*, **17**(2–3), 295–324, (1997).

[11] E. Hisdal, 'Conditional possibilities independence and non interaction', *Fuzzy Sets and Systems*, 283–297, (1978).

[12] R. C. Jeffrey, *The Logic of Decision*, McGraw Hill, NY, 1965.

[13] F. V. Jensen and Thomas D. Nielsen, *Bayesian Networks and Decision Graphs (Information Science and Statistics)*, Springer, June 2007.

[14] O. Papini, 'Iterated revision operations stemming from the history of an agent's observations.', in *In H. Rott and M. Williams, editors, Frontiers in Belief Revision, pages 279-301. Kluwer Academic Publishers, 2001.*

[15] J. Pearl, *Probabilistic reasoning in intelligent systems: networks of plausible inference*, Morgan Kaufmman, San Francisco, 1988.

[16] J. Pearl, *Causality: Models, Reasoning, and Inference*, Cambridge University Press, March 2000.