Possibilistic logic bases and possibilistic graphs

Salem BENFERHAT - Didier DUBOIS - Laurent GARCIA - Henri PRADE

Institut de Recherche en Informatique de Toulouse (I.R.I.T.)
Université Paul Sabatier — 118, route de Narbonne
31062 Toulouse — France
{benferhat, dubois, garcia, prade}@irit.fr

Abstract

Possibilistic logic bases and possibilistic graphs are two different frameworks of interest for representing knowledge. The former stratifies the pieces of knowledge (expressed by logical formulas) according to their level of certainty, while the latter exhibits relationships between variables. The two types of representations are semantically equivalent when they lead to the same possibility distribution (which rank-orders the possible interpretations). A possibility distribution can be decomposed using a chain rule which may be based on two different kinds of conditioning which exist in possibility theory (one based on product in a numerical setting, one based on minimum operation in a qualitative setting). These two types of conditioning induce two kinds of possibilistic graphs. In both cases, a translation of these graphs into possibilistic bases is provided. The converse translation from a possibilistic knowledge base into a min-based graph is also described.

1 Introduction

Possibilistic logic is issued from Zadeh's possibility theory [16], which offers a framework for the representation of states of partial ignorance owing to the use of a dual pair of possibility and necessity measures. Possibility theory may be quantitative or qualitative according to the range of these measures which may be the real interval [0,1] or a finite linearly ordered scale as well [8]. Possibilistic logic (e.g., [7]) has been developed for more than ten years. It provides a sound and complete machinery for handling qualitative uncertainty with respect to a semantics expressed by means of possibility distributions which rank-orders the possible interpretations. At the syntactic level, possibilistic logic handles pairs of the form (p α) where p is a classical logic formula and α is an element of a totally ordered set. The pair (p α) expresses that the formula p is certain at least to the level α, or more formally by $N(p) \geq \alpha$, where N is the necessity measure associated to the possibility distribution expressing the underlying semantics. Possibilistic logic is essentially *qualitative* since only the preordering induced on the formulas is important (N(p) > N(q) means "p is more certain than q"). Possibilistic logic

is as tractable as classical logic since its complexity is about $\log_2 n*SAT$ where n is the number of certainty levels used in the knowledge base and SAT is the complexity of satisfaction problem in classical logic. Besides, there exist few works on directed possibilistic graphs (which are the counterpart of Bayesian probabilistic networks [13,14] in the framework of possibility theory). Existing works are either a direct adaptation of a probabilistic logic approach without caring for knowledge representation [10], or a way to do learning from imprecise data [11]. Because of the existence of the possibilistic logic machinery, there is not the same necessity to introduce graphical structures in possibility theory, as in probability theory (where probabilistic logic is more complex to handle). Yet Bayesian-like networks have a clear appeal for knowledge acquisition and directed graphs could be used to help in the specification of possibilistic knowledge as much as probabilistic knowledge.

Next section gives a background on possibilistic logic, on conditioning in possibility theory and on the directed possibilistic graphs. Section 3 studies their encoding in possibilistic logic. Section 4 discusses the problem of recovering the initial conditional possibility distribution from the joint possibility distribution computed with the chain rule, and the discussion briefly refers to the idea of possibilistic independence. Section 5 proposes an encoding of a set of possibilistic logic formulas into directed possibilistic graphs. Proofs are omitted for the sake of brevity, but can be found in [2].

2 Background

2.1. Possibilistic knowledge bases

A possibilistic (knowledge or belief) base is a set of possibilistic logic formulas of the form (p,α) where p is a classical propositional logic formula; α an element of the semi-open real interval (0,1] in a numerical setting, or of a finite linearly ordered scale in a qualitative setting. It estimates to what extent it is certain that p is true considering the available, possibly incomplete information about the world.

Given a possibilistic base Σ, we can generate a possibility distribution from Σ by associating to each classical interpretation a degree in [0,1] expressing the level of compatibility with the available information. When a

possibilistic base is made of one formula $\{(p\ \alpha)\}$, then each interpretation ω which satisfies p gets the degree $\pi(\omega) = 1$ since it is completely consistent with p, and each interpretation ω which falsifies p gets a degree $\pi(\omega)$ such that the highest is α (i.e., the more certain is p), the lowest is $\pi(\omega)$. In particular, if $\alpha=1$ (i.e., p is completely certain), then $\pi(\omega) = 0$, i.e., ω is impossible. One way to realize this constraint is to assign to $\pi(\omega)$ the degree $1 - \alpha$ (on an ordered scale, we use a reversing map of the scale). Then, the possibility distribution associated to $\{(p\ \alpha)\}$ is:

$$\forall \omega \in \Omega,\ \pi_{\{(p\ \alpha)\}}(\omega) = 1 \quad \text{if } \omega \models p$$
$$= 1 - \alpha \quad \text{otherwise.}$$

When $\Sigma = \{(p_i, \alpha_i), i=1,n\}$ is a general possibilistic base then all the interpretations satisfying all the beliefs in Σ will have the highest possibility degree, namely 1, and the other interpretations are ranked w.r.t. the highest belief that they falsified, namely we get $\forall \omega \in \Omega$:

$$\pi_\Sigma(\omega) = 1 \quad \text{if } \omega \models \Sigma$$
$$= 1 - \max\{\alpha_i : (p_i\ \alpha_i) \in \Sigma \text{ and } \omega \models \neg p_i\} \quad \text{otherwise.}$$

Thus, π_Σ can be viewed as the result of the combination of the $\pi_{\{(p_i\ \alpha_i)\}}$'s using the min operator, i.e.:

$$\pi_\Sigma(\omega) = \min\{\pi_{\{(p_i\ \alpha i)\}}(\omega) : (p_i\ \alpha_i) \in \Sigma\}.$$

A possibility distribution π_Σ is said to be *normal* if there exists an interpretation ω which is totally possible, namely $\pi_\Sigma(\omega) = 1$. However, in general there may exist several distinct interpretations which are totally possible. This *normalization condition* reflects the consistency of the available knowledge Σ represented by this possibility distribution (i.e., $\exists \omega \in \Omega$, s.t., $\omega \models \Sigma$). A possibility distribution π induces two mappings grading respectively the possibility and the certainty of a formula p:

– the possibility degree $\Pi(p) = \max\{\pi(\omega) : \omega \models p\}$ which evaluates to what extent p is consistent with the available knowledge expressed by π. Note that we have:

$$\forall p\ \forall q \quad \Pi(p \vee q) = \max(\Pi(p), \Pi(q));$$

– the necessity (or certainty) degree $N(p) = \min\{1 - \pi(\omega) : \omega \models \neg p\}$ which evaluates to what extent p is entailed by the available knowledge. We have:

$$\forall p\ \forall q \quad N(p \wedge q) = \min(N(p), N(q)).$$

It can be checked that if $(p,\alpha) \in \Sigma$, the semantic constraint $N(p) \geq \alpha$ holds, where N is a necessity measure based on π_Σ. Note the duality equation: $N(p) = 1 - \Pi(\neg p)$. Moreover, note that, contrasting with probability theory $N(p)$ and $N(\neg p)$ (resp. $\Pi(p)$ and $\Pi(\neg p)$) are not functionally related: we only have (for normal possibility distributions) $\min(N(p), N(\neg p)) = 0$ (resp. $\max(\Pi(p), \Pi(\neg p)) = 1$). It leaves room for representing complete ignorance in an unbiased way: p is ignored whenever $\Pi(p) = \Pi(\neg p)) = 1$.

Lastly, several syntactically different possibilistic belief bases may have the same possibility distribution as a semantic counterpart. In such a case, it can be shown that these bases are equivalent in the following sense: their α-cuts, which are classical bases, are logically equivalent in the usual sense, where the α-cut of a possibilistic base Σ

is the set of classical formulas whose level of certainty is greater than or equal to α.

2.2. Possibilistic graph

Let $V = \{A_1,...,A_n\}$ be a set of variables. These variables will identify the nodes of a network. In possibilistic logic, these variables are binary. The domains associated with the variables A_i are denoted by $D_V=\{D_1,...,D_n\}$. A possible assignment of a value of D_i to every variable A_i will be called an elementary event. When each variable is binary (i.e., $D_i=\{a_i,\neg a_i\}$), the elementary events are also called interpretations and are denoted by ω. The set of all the elementary events is simply the Cartesian product $D_1\times...\times D_n$ of the domains. Formulas, represented as sets of interpretations, are also called events.

Possibilistic graphs, denoted by ΠG, as probabilistic networks, are based on directed acyclic graph [10, 11]. The nodes represent variables (for example, the temperature of a patient, the colour of a car, ...) and the edges encode the causal link (or influence) between these variables. Uncertainty is represented on each node. When there is an edge from the node A_i to the node A_j, the node A_i is said to be "parent" of A_j (parents of A are denoted by $Par(A)$). A possibility measure Π is associated with a graph G in the following way:

• For the nodes A_i which are roots of the graph (i.e., $Par(A_i)=\varnothing$), we specify the prior possibility degrees associated with each instance of A_i, namely we give every $\Pi(a)$ where $a \in D_i$. Possibilities must satisfy the normalization condition: $\max_{a \in D_i} \Pi(a)=1$.

• To each other node A_j are attached conditional possibilities $\Pi(a \mid \omega_{Par(Aj)})$ where $a \in D_j$ et $\omega_{Par(Aj)}$ is an element of the Cartesian product $X_i D_i$ of the $D_i \in Par(A_j)$, domains associated with the variables of $Par(A_j)$. Conditional possibilities should satisfy the following normalization condition:

$$\forall \omega_{Par(A)} \in X_i D_i,\ \max_{a \in D_j} \Pi(a \mid \omega_{Par(A)}) = 1,$$

$\Pi(.|b)$ is a conditional possibility measure, which can be defined in two different ways (as explained after the following Definition). Then:

Definition 1: *The joint possibility distribution associated with a possibilistic graph ΠG is computed with the following equation:*

$$\pi^\square_{\Pi G}(x_1...x_n) = \square_{i=1,n}\ \Pi(x_i \mid \omega_{x_i}) \quad (1),$$

where $x_i \in \{a_i, \neg a_i\}$ are the two possible instances of the variable A_i, $\omega_{x_i} \subseteq \{x_1...x_n\}$ and $\square$ represents either the minimum or the product. We shall use π_* and π_m for short, in a case $\square$ = product or $\square$ = min respectively.

Indeed, in possibility theory, there exist two definitions for the conditioning:

• In a qualitative setting the conditioned possibility measure is defined by:

$$\Pi(q|p) = 1 \quad \text{if } \Pi(p \wedge q) = \Pi(p)$$
$$= \Pi(p \wedge q) \text{ otherwise (i.e., } \Pi(p \wedge q) < \Pi(p)),$$

and obeys the following equation [12] :

$$\Pi(p \wedge q) = \min(\Pi(q|p), \Pi(p)); \quad (2)$$

This definition of conditioning only requires the ordering between interpretations, and can be defined on any finite ordeed scale.

• In a numerical setting we use the following definition of conditioning based on the product :

$$\Pi(p \wedge q) = \Pi(q|p) * \Pi(p) \quad \text{if } \Pi(p) \neq 0$$
$$= 0 \qquad \text{otherwise.} \qquad (3)$$

In both cases we have $N(q|p) = 1 - \Pi(\neg q|p)$. Up to a rescaling, (3) is also the conditioning rule of kappa-functions [15]:

$$\kappa(\omega|p) = \kappa(\omega) - \kappa(p) \quad \text{if } \omega \models p$$
$$= \infty \qquad \text{otherwise,}$$

with $\Pi(\phi) = 2^{-\kappa(\phi)}$. Conditioning a possibility distribution with p then with r gives the same result as conditioning with r and then with p.

When the joint possibility distribution (1) is computed with the minimum (resp. product), the ΠG will be denoted by ΠG_{min} (resp.$\Pi G*$).

2.3. Decomposition

The decomposition of a possibility distribution consists in expressing a joint possibility distribution as a combination of conditional possibility distributions. For this decomposition, we can follow the same way as in probability theory. Let $\{A_1,...,A_n\}$ be the set of variables which is ordered arbitrarily. From the definition of conditioning, we have :

$$\pi(A_1...A_n) = \min [\pi(A_n|A_1...A_{n-1}), \pi(A_1...A_{n-1})].$$

Applying repeatedly this definition to $\pi(A_1...A_{n-1})$, then to $\pi(A_1...A_{n-2})$ ..., the joint possibility distribution is decomposed into

$$\pi(A_1...A_n) = \min [\pi(A_n|A_1...A_{n-1}) ,..., \pi(A_1)]. \qquad (4)$$

The decomposition given by equation (4) can be simplified by assuming conditional independence between variables. When conditioning with the product, the decomposition follows the way used in probability:

$$\pi(A_1...A_n) = \pi(A_n|A_1...A_{n-1})*...*\pi(A_1). \qquad (5)$$

There is no unique decomposition of a possibility distribution since it depends on the initial ordering between variables as it is shown in the following example:

Example 1: Let π be defined on $\{a, \neg a\} \times \{b, \neg b\}$:
$\pi(\neg a \neg b) = 1$; $\pi(a\ b) = \pi(a\ \neg b) = .8$; $\pi(\neg a\ b) = .7$.
There are two ways for decomposing this possibility distribution,

• either $\pi(A, B) = \min (\pi(B|A), \pi(A))$, and hence $\pi(B|A)$ is given by the matrix:

B|A	a	¬a
b	1	.7
¬b	1	1

and $\pi(A)$ by the values $\pi(a) = .8$ and $\pi(\neg a) = 1$,

• or $\pi(A, B) = \min (\pi(A|B), \pi(B))$, and hence $\pi(A|B)$ is given by the matrix:

A|B	b	¬b
a	1	.8
¬a	.7	1

and $\pi(B)$ by the values $\pi(b) = .8$ and $\pi(\neg b) = 1$.

It can be verified that the two possible decompositions of the possibility distribution π correspond to the two following possibilistic knowledge bases:

$$\Sigma_1 = \{(\neg b \vee a\ .3), (\neg a\ .2)\},$$
$$\Sigma_2 = \{(b \vee \neg a\ .2), (\neg b \vee a\ .3), (\neg b\ .2)\}.$$

These two bases (which can be computed as explained in next sections) are equivalent since they are associated with the same possibility distribution (i.e., $\pi_{\Sigma 1} = \pi_{\Sigma 2}$).

3. From the graph to the logical base

The goal of this section is to translate a directed possibilistic graph into a possibilistic base. The translation should be such that the possibility distribution associated to the graph using the chain rule is the same as the one associated to the possibilistic knowledge base. The restriction to binary variables is made for the sake of simplicity. See [1] for an extension to the non-binary case.

The directed possibilistic graph can be seen as the result of the fusion of one-formula knowledge bases, each one containing a single possibilistic formula. Each formula corresponds to a conditional possibility of the directed possibilistic graph. A possibilistic causal network is viewed as a set of triples:

$$\Pi G = \{(a, P_a, \alpha) : \Pi(a|P_a) = \alpha \neq 1 \text{ is an element of the directed graph}\}$$

where 'a' is an instance of the variable A and P_a is an element of the Cartesian product $X_i D_i$ of the D_i. It can be restricted to the conditional possibilities different from 1 since the ones which are equal to 1 are not used in the computation of joint possibility distributions.

3.1   Encoding ΠG_{min}

Let us start with min-based conditioning. With each triple (a, P_a, α) of the directed possibilistic graph is associated the single formula $(\neg a \vee \neg P_a\ 1-\alpha)$. The conditional symbol $\neg a|P_a$ is encoded as the material implication $\neg a \vee \neg P_a$, using the duality $\Pi(a|P_a) = 1 - N(\neg a|P_a)$.

The joint possibility distribution obtained from a directed possibilistic graph using (1) is equivalent to the one obtained by min-combination of the possibility distributions associated with the formulas encoding the different triples of the directed graph:

Proposition 1: Let π_{a_i} be the possibility distribution associated with the formula corresponding to the triple (a_i, P_{a_i}, α_i). Then: $\pi_m = \min_i \pi_{a_i}$.

As already suggested in the background section, the union of two possibilistic bases Σ_1 and Σ_2 corresponds to the min-combination of the two possibility distributions π_1 and π_2 associated with Σ_1 and Σ_2 respectively. Therefore, the following Proposition states the knowledge base associated with a directed possibilistic graph:

Proposition 2: The possibilistic knowledge base associated with ΠG_{min} is:
$$\Sigma = \{(\neg y_i \vee \neg P_{y_i}\ 1-\alpha_i)) / (y_i, P_{y_i}, \alpha_i) \in \Pi G\}.$$

This result is important since it implies that every result known for possibilistic logic can be applied to directed possibilistic graphs.

Example 2 : Consider the following DAG :
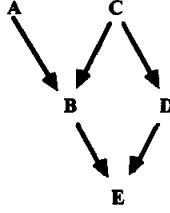
Assume that the conditional possibility degrees are given by the following tables :

$\Pi(A)$

a	1
¬a	3/4

$\Pi(C)$

c	1
¬c	1

$\Pi(B \mid A\,C)$

B | A C	a c	¬a c	elsewhere
b	1/2	1	1
¬b	1	1/4	1

$\Pi(D \mid C)$

D | C	c	¬c
d	1	1/4
¬d	1	1

$\Pi(E \mid B\,D)$

E | B D	b ¬d	¬b d	elsewhere
e	1	1/2	1
¬e	3/4	1	1

We have:
$R\Pi$ = {(¬a, $\varnothing$, 3/4), (b, a c, 1/2), (¬b, ¬a c, 1/4), (d, ¬c, 1/4), (¬e, b ¬d, 3/4), (e, ¬b d, 1/2)}.
Then : Σ = {(a 1/4), (¬bv¬av¬c 1/2), (bvav¬c 3/4), (¬dvc 3/4), (ev¬bvd 1/4), (¬evbv¬d 1/2)}.

3.2. Encoding ΠG^*

Let us now turn towards product-based conditioning. Following the same approach as above, to each triple (a, P_a, α) of a possibilistic graph, is associated the single possibilistic formula (¬av¬P_a 1–α). Let us notice that:
$$\Pi_a(a \mid P_a) = \Pi_a(a \wedge P_a) / \Pi_a(P_a) = \alpha,$$
where Π_a is the possibility measure obtained from the possibility distribution associated with (¬av¬P_a 1–α). Indeed, $\Pi_a(a \wedge P_a) = \alpha$ since each interpretation satisfying $a \wedge P_a$ falsifies (¬av¬P_a 1–α). Moreover, $\Pi_a(P_a) = \max(\Pi_a(a \wedge P_a), \Pi_a(\neg a \wedge P_a))$. As every interpretation satisfying ¬a$\wedge P_a$ satisfies (¬av¬P_a 1–α), we get $\Pi_a(P_a)=1$. The counterpart of Proposition 1 holds:

Proposition 3: Let π_{a_i} be the possibility distribution associated with (a_i, P_{a_i}, α_i). Then $\pi_* = *_i\, \pi_{a_i}$, where $*$ is the pointwise product of possibility distributions.

We need to find what is the possibilistic base associated with the combination of two possibility distributions by the product operator. This is summarized in the following proposition [3]:

Proposition 4: Let Σ_1={($p_i\ \alpha_i$) : $i \in I$} and Σ_2={($q_j\ \beta_j$) : $j \in J$}. Let π_1 and π_2 be the two possibilistic distributions associated with Σ_1 and Σ_2. Let π^* be the product-based combination of π_1 and π_2. The resulting base associated with π^* is:
$$\check{C}^*(\Sigma_1, \Sigma_2) = \Sigma_1 \cup \Sigma_2 \cup$$
{($p_i \vee q_j$ ($\alpha_i + \beta_j - \alpha_i*\beta_j$)) : ($p_i\ \alpha_i) \in \Sigma_1$ and ($q_j\ \beta_j) \in \Sigma_2$}.

Since the operator $\check{C}^*$ is commutative and associative, this straithforwardly extends to the case of more than two bases. Hence, the possibilistic base which is associated to ΠG^* can be obtained by combining the different triples of the graph with the syntactic operator $\check{C}^*$.

Example 2 (continued) : Recall that :
$R\Pi$={(¬a, $\varnothing$, 3/4), (b, a c, 1/2), (¬b, ¬a c, 1/4), (d, ¬c, 1/4), (¬e, b ¬d, 3/4), (e,¬b d,1/2)}.
The six elementary knowledge bases associated to the triples of ΠG^* are:

Σ_1={(a 1/4)} ; Σ_2={(¬av¬bv¬c 1/2)} ;
Σ_3={(avbv¬c 3/4)} ; Σ_4={(cv¬d 3/4)} ;
Σ_5={(¬bvdve 1/4)} ; Σ_6={(bv¬dv¬e 1/2)}.

Now let us combine them with $\check{C}^*$:
- the combination of Σ_1 et Σ_2 leads to:
  $\Sigma_{12}=\check{C}^*(\Sigma_1, \Sigma_2)$ = {(a 1/4), (¬av¬bv¬c 1/2)} ;

- combining the result with Σ_3 :
  $\Sigma_{123}=\check{C}^*(\Sigma_{12}, \Sigma_3)$
  = {(a 1/4), (¬av¬bv¬c 1/2), (avbv¬c 13/16)} ;

- combining the result with Σ_4 :
$\Sigma_{1234}=\check{C}^*(\Sigma_{123}, \Sigma_4)$
= {(a 1/4), (¬av¬bv¬c 1/2), (avbv¬c 13/16), (cv¬d 3/4), (avcv¬d 13/16)} ;

- combining the result with Σ_5 :
$\Sigma_{12345}=\check{C}^*(\Sigma_{1234}, \Sigma_5)$
= {(a 1/4), (¬av¬bv¬c 1/2), (avbv¬c 13/16), (cv¬d 3/4), (avcv¬d 13/16), (¬bvdve 1/4), (av¬bvdve 7/16), (¬av¬bv¬cvdve 5/8)} ;

- combining the result with Σ_6 leads to the final base :
$\Sigma^*=\check{C}^*(\Sigma_{12345}, \Sigma_6)$
= {(a 1/4), (¬av¬bv¬c 1/2), (avbv¬c 13/16), (cv¬d 3/4), (avcv¬d 13/16), (¬bvdve 1/4), (av¬bvdve 7/16), (¬av¬bv¬cvdve 5/8)} $\cup$ {(bv¬dv¬e 1/2)} $\cup$ {(avbv¬dv¬e 5/8), (avbv¬cv¬dv¬e 29/32), (bvcv¬dv¬e 7/8), (avbvcv¬dv¬e 29/32)}.

This knowledge base contains 13 clauses while in the case of min we only have 6. This clearly illustrates that

the combination with the product leads to a larger knowledge base than if we combined with the minimum, due to Proposition 4. This comes with the fact that the product is not compatible with a finite scale.

4. Recovering initial data

A natural question when we compute the joint possibility distribution using the chain rule is to see if we recover the a priori and conditional probabilities given by the expert. In the probability theory the answer is always yes. The following proposition shows that this is also the case if the chaining rule is based on the product:

Proposition 5: Let $\Pi(a|Pa)$ be the conditional possibility distributions over the variables A in the $\Pi G*$. Let $\pi*$ be the joint possibility distribution obtained using the chain rule with the product (1), and its associated possibility measure $\Pi*$. Then, for any conditional possibility distribution we have:

$$\Pi*(a|Pa) = \Pi(a|Pa).$$

Concerning the minimum-based conditioning the answer is no, as it is illustrated by the following small example:

Example 3 Let us consider the following graph:

with $\Pi(a)=1$; $\Pi(\neg a)=1/4$; $\Pi(b|a) = 1/3$; $\Pi(\neg b|a) = 1$ and $\Pi(b|\neg a) = 1/3$; $\Pi(\neg b|\neg a) =1$.
The joint possibility distribution is:

$\pi_m(ab) = 1/3$; $\pi_m(a\neg b) = 1$;
$\pi_m(\neg ab) = 1/4$; $\pi_m(\neg a\neg b) = 1/4$;

Clearly, we have $\Pi_m(b|\neg a) = 1 \neq 1$ since $\pi_m(\neg ab) = \pi_m(\neg a) = 1/4$.

The reason for not recovering the original values is that the conditional possibility distributions specified by the users are not coherent with the axioms of possibility distributions. Indeed, using the definition of conditional possibility measure, we always have:

If $\Pi(p|q) \neq 1$ then $\Pi(p|q) = \Pi(p \wedge q) < \Pi(q)$.

We see clearly, from the previous example that this constraint is violated since $\Pi(b|\neg a) = 1/3 \neq 1$ and $\Pi(b|\neg a) > \Pi(\neg a)=1/4$. Therefore, it is not surprizing if we do not recover the above value.
This behaviour also exists in possibilistic logic, namely a possibility distribution associated a possibilistic base do not guarantee to recover the exact value of the knowledge base. To be convinced, just consider a small example where $\Sigma = \{(a \ .8), (a \vee b \ .4)\}$. We can easily check that $N_{\pi\Sigma}(a \vee b) = .8$. This is due to the fact that $(a \vee b \ .4)$ is strictly subsumed by $(a \ .4)$. Hence, we have:

Proposition 6: Let Σ be such that it does not contain any strictly subsumed formulas, namely there is no (p, α) such that $\{q : (q, \beta) \in \Sigma$ and $\beta > \alpha\} \vdash p$. Then $\forall (p, \alpha) \in \Sigma$ we have: $N_{\pi\Sigma}(p) = \alpha$.

Let us go back to the causal network, the following proposition characterizes to what the computed joint possibility corresponds w.r.t. the conditional possibility distributions specified by the user.

Proposition 7: Let $\Pi(a|Pa)$ be the conditional possibility distributions over the variables A in the DAG. Let π_m be the joint possibility distribution obtained using the chain rule with the minimum-based conditioning. Then: either $\pi_m(a|Pa) = \Pi(a|Pa)$ or $\pi_m(a|Pa) = 1$.

This means that the computed joint possibility distribution either preserves the initial values or push them up to 1 (this is observed in Example 1).
We now focus on the DAG, from where we recover the original values. In the possibility theory, the definition of independence is not unique [4][6]. The following is a weaker one called "non-interactivity":

Definition 2: Two variables A and B are independent in the context C, if for each instance (a,b,c) of (A, B, C) we have: $\Pi(a,b \ |c) = \min (\Pi(a|c), \Pi(b|c))$.

The following proposition shows that the joint possibility distribution guarantees the independence relations from the structure of the DAG, as in probabilistic network:

Proposition 8: Let X be a given variable, and Y be a variable which is neither a parent of X nor any of its descendant. Let π_m be the joint possibility distribution computed from a DAG ΠG using the min-based chain rule. Then X and Y are independent in the context of Par(X) in the sense of Definition 2.

The question of whether a joint possibility distribution can be decomposed using a stronger definition of independence is left for further research.

5. Encoding bases into min-based graphs

In this section, we present the transformation of possibilistic knowledge bases into directed possibilistic graphs ΠG_{min}. One way to do it is to use possibility distributions as intermediary. Indeed, a knowledge base leads to a possibility distribution, from which it is possible to build a ΠG_{min} (see Section 2.3). This would apply as well to $\Pi G*$. However, this way is computationally expensive. Moreover, we want to find the ΠG_{min} directly from the knowledge base.
The encoding of a possibilistic knowledge into a ΠG_{min} is less straightforward than the previous transformations. Indeed, we cannot directly view each formula as a triple and then build the graph, but we need some pre-processing steps. The constructed possibilistic graph ΠG should be such that:
• the joint possibility distribution computed from the ΠG using the minimum operator should be the same as the one computed from the knowledge base;

- the joint possibility distribution allows us to recover all the conditional possibilities of the DAG, and
- the joint possibility distribution satisfies the independence relation (in the sense of Definition 2) induced by the structure of the DAG, namely each variable should be independent of any, non-descendent, variable in the context of its parents.

The construction of a causal network associated to a possibilistic knowledge is obtained in three steps: the first step simply consists in putting the knowledge base into a clausal form and in removing tautologies. The second step consists in constructing the graph, and the last step computes the conditional possibilities associated to the constructed graph.

5.1 Putting bases in a clausal form

In this first step, a base Σ is rewritten into an equivalent base Σ'. The equivalence means here: $\pi_\Sigma = \pi_{\Sigma'}$. Getting Σ' consists in putting the knowledge base into a clausal form and in removing tautologies. The following proposition shows how to put the base in a clausal form first:

Proposition 9 : Let $(p \ \alpha) \in \Sigma$. Let $\{c_1, ..., c_n\}$ be the set of clauses encoding p in classical logic. Let Σ^* be a new knowledge base obtained from Σ by replacing $(p \ \alpha)$ by $\{(c_1 \ \alpha), ..., (c_n \ \alpha)\}$. Then the two knowledge bases Σ and Σ^* are equivalent in the sense that $\pi_\Sigma = \pi_{\Sigma^*}$.

Then removing tautologies leads still to an equivalent possibilistic base. Indeed, tautologies are always satisfied, and $\pi_\Sigma(\omega)$ is only defined with respect to that formulas of Σ falsified by ω. The removing of tautologies is an important point since this will avoid having links which do not make sense. For example, the tautological formula $(\neg x \lor \neg y \lor x \ 1)$ could induce a link between X and Y.

5.2 Constructing the graph

The second step consists in constructing the graph, namely the determination of the vertices (variables) of the graph and the parents of each vertex. The set of variables is simply the set of propositional symbols which appear in the knowledge base. Moreover, since possibilistic logic is based on propositional logic, then all the used variables are binary variables. By X we denote the variable which can be either x or $\neg x$. To construct the graph, we first rank the variables, according to an arbitrary numbering $\{X_1, X_2, ..., X_n\}$ of the variables. This ranking intends to mean that parents of each variable X_1 can only be in $\{X_{i+1}, ..., X_n\}$ (but they may not exist).
We first give several intuitive examples before presenting the technical construction of the graph.

Example 4: Let: $\Sigma = \{(t \ 0.6), (t \lor v \ 0.4)\}$. From this knowledge base one may think that the variable T depends on the variable V. However, we can easily check that Σ is equivalent to the following one: $\Sigma' = \{(t \ 0.6))\}$, where clearly, V has no influence on T. The formula $(t \lor v \ 0.4)$ is said to be subsumed by $\Sigma - \{(t \lor v \ 0.4)\}$.

Next definition formally introduces the notion of subsumed beliefs, which can be removed as stated by Proposition 10:

Definition 3: A formula $(p \ \alpha)$ of Σ is said to be subsumed if $\Sigma_{\geq \alpha} \models p$ where $\Sigma_{\geq \alpha} = \{q : (q \ \beta) \in \Sigma,$ and $\beta \geq \alpha\} - \{(p \ \alpha)\}$ (namely p can be recovered from clauses of the base having weights greater or equal than α).

Proposition 10: Let Σ' be a new base obtained from Σ by removing subsumed formulas. Then $\pi_\Sigma = \pi_{\Sigma'}$.

Therefore, we can remove, or add, subsumed beliefs without any damage. Subsumed beliefs are not the only ones which may induce fictitious dependencies:

Example 5: Let: $\Sigma = \{(a \ .5), (\neg a \lor b \ .5)\}$.
In this base, neither $(a \ .5)$ nor $(\neg a \lor b \ .5)$ is subsumed, and one can think that there is a relationship between the variables A and B. However, we can easily check that this base is equivalent to $\Sigma' = \{(a \ .5), (b \ .5)\}$, where clearly A and B are unrelated. So we are let to state:

Proposition 11: Let X be a variable, and $(x \lor p \ \alpha)$ be a clause of Σ containing the instance x of X s.t. $\Sigma \vdash (p \ \alpha)^1$. Then the base Σ and the knowledge base Σ' obtained from Σ by replacing $(x \lor p \ \alpha)$ by $(p \ \alpha)$ are equivalent.

Intuitively, one can say that two variables are related if there is a clause containing an instance of these two variables, and they are unrelated otherwise. Example 6 shows that two variables can be related even if there is no clause in the base containing an instance of variables:

Example 6: Let: $\Sigma = \{(x \lor a \ .5), (\neg x \lor b \ .5)\}$.
In this base, if we would define the relationship between variables only if there exists a clause containing an instance of each of them, then clearly A and B would be unrelated. However, we can check that: $\Sigma \vdash (a \lor b \ .5)$.

The following example shows that in order to compute the conditional possibility distribution of a variable given its parents, it is not enough to look only to clauses of the base containing instances of the variable X:

Example 7: Let: $\Sigma = \{(a \lor b \ .6), (x \lor b \ .5), (x \lor a \ .5)\}$.
We assume that parents of X are A and B. Clearly, if we compute the conditional possibilities $\Pi(X|AB)$ only from clauses containing X, namely $\Sigma_X = \{(x \lor b \ .5), (x \lor a \ .5)\}$ then it is not guaranteed to recover all original values. Indeed, in this example, if the computation of $\Pi(\neg x | \neg a \neg b)$ is simply based on Σ_X we get: $\Pi(\neg x | \neg a \neg b) = .5$
$\quad$ (since $\{(x \lor b \ .5), (x \lor a \ .5)\} \vdash (x \lor a \lor b \ .5)$)
but we can check that after computing the joint possibility distribution π_Σ: $\Pi_\Sigma(\neg x | \neg a \neg b) = 1$, this is due to the fact that we have both: $\Sigma \vdash (x \lor a \lor b \ .6)$ and $\Sigma \vdash (a \lor b \ .6)$, hence: $\Pi_\Sigma(\neg x \land \neg a \neg b) = \Pi_\Sigma(\neg a \neg b)$.
Indeed, $(x \lor b \ .5)$ and $(x \lor a \ .5)$ are not subsumed but $(x \lor b \lor b$

----

^{1} This is the syntactic possibilistic entailment. We have:
$\Sigma \vdash (p, \alpha)$ iff $\{q : (q, \beta) \in \Sigma$ and $\beta \geq \alpha\} \cup \{\neg p\}$ is inconsistent.

.5) is because of the clause (a∨b .6) in Σ.

The following definition enables a direct computation of conditional possibility distributions:

Definition 4: Let X be a variable and Par (X) be the set of its parents. Let K be a set of clauses of the form x∨p such that P⊆Par (X) (P is the set of variables containing an instance in x∨p). We call complete extension of K, denoted by E(K), the set of all clauses of the form $(x \vee \bigvee_{y \in P_X} y, \alpha)$ where x is an instance of X, P_X is an instance of Par (X), and

$\alpha = \max \{a_i : (x \vee p_i \ a_i) \in K \text{ and } p_i \models \bigvee_{y \in P_X} y\}$,
with max $\{\varnothing\}$=0.

Example 8: Let K = {(x∨b .5), (x∨a .5)}. Assume that the parents of X are A and B. Then we can check that: E(K) = {(x∨b∨¬a .5), (x∨b∨a .5), (x∨a∨¬b .5)}.

Proposition 12: The two bases K and E(K) are equivalent.

With the help of the previous propositions we are now ready to determine the exact parents of each variable. Let $\{X_1, ..., X_n\}$ be a numbering of the variables. Then the following algorithm determines the set of parents of each variable:

For i = 1,..., n do
Begin
    /* Determination of Parents of X_i */
1. Let $(x_i \vee p, \alpha)$ be a clause of Σ s. t. :
    x_i is an instance of X_i, and
    p is only built from $\{X_{i+1},...,X_n\}$.
   • If $(x_i \vee p, \alpha)$ is subsumed, then remove it from Σ
     (due to Prop. 10).
   • If $\Sigma \vdash (p, \alpha)$ then replace $(x_i \vee p, \alpha)$ by (p, α)
     (due to Prop. 11)
2. Let K_i be the set of clauses $(x_i \vee p, \alpha)$ in Σ s. t.
    p is only built from $\{X_{i+1},...,X_n\}$
3. The parents of the variable X_i are :
    Par(X_i)= $\{X_j : \exists \ c \in K_i$ such that c contains
        an instance of $X_j\}$
    /* Rewriting Σ to facilitate the computation of
    conditional possibilities , and for looking for hidden
    dependencies */
4. Replace in Σ, K_i by its complete extensions
    $E(K_i)$ (due to Prop. 12).
5. For each $(x_i \vee p \ \alpha)$ of Σ (where p is built from
    $\{X_{i+1},...,X_n\}$) such that $\Sigma \vdash (p, \alpha)$
        replace $(x_i \vee p, \alpha)$ by (p, α).
6. Let Σ_{Xi} be the set of clauses $(x_i \vee p, \alpha)$ in Σ s. t.
    p is only built from $\{X_{i+1},...,X_n\}$
End.

The algorithm is composed of two parts: the first part (Steps 1-3) consists in the construction of the DAG by determining the parents of each variable, and the second part (Steps 4-6) consists in rewriting the knowledge base

such that : i) it gives immediately the conditional possibility distributions, and ii) ensures to recover original values when using the chain rule for computing the joint possibility distributions. Indeed, once $E(K_i)$ is computed, in order to evaluate $\Pi(x|p_X)$ then either $(\neg x_i \vee \neg p, \alpha) \notin E(K_{X_i})$ then $\Pi(x|p_X) = 1$, or $(\neg x_i \vee \neg p, \alpha) \in E(K_i)$, then if $\Sigma \vdash (\neg p, \alpha)$ then $\Pi(x|p_X) = 1$ (since $\Pi(x \wedge p_X) = \Pi(x)$), otherwise $\Pi(x|p_X) = 1 - \alpha$.

The result of the algorithm is a partition $\{\Sigma_{X1}, ..., \Sigma_{Xn}\}$ such that $\Sigma_{X1} \cup ... \cup \Sigma_{Xn}$ is equivalent to Σ. Clearly, for i>1, Σ_{Xi} does not contain any variable of $\{X_1, ..., X_{i-1}\}$. Moreover, Σ_{Xi} can be empty. In this case X_i has no parents : it corresponds to the root of the graph, and the priori possibility degrees associated to X_i are equal to 1.
The subbases Σ_{Xi}'s give a direct computation of conditional possibility degrees as explained later.

A graph associated to Σ is such that its vertices are the variables in Σ, and a link is drawn from Xj to Xi iff Xj ∈ Par(X_i), where Par(X_i) is given by step 3 in the algorithm. This graph is indeed a DAG.
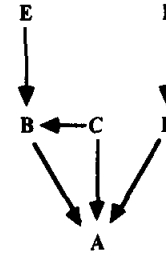
Example 9 : Let us consider the following base: Σ = {(a∨b 0.7), (¬a∨c∨¬d 0.7), (a∨c∨d 0.9), (b∨c 0.8), (¬b∨e 0.2), (¬d∨f 0.5)}
Σ contains six variables numbered in the following way: X_1=A, X_2=B, X_3=C, X_4=D, X_5=E et X_6=F.
From this partition we can check that we get:
Par(A)={B, C, D}, Par(B)={C, E}, Par(C)=∅,
Par(D)={F}, Par(E)= PAR(F)=∅. The final graph is:

5.3 Determining the conditional possibilities
Once the graph is constructed we need to compute the conditional possibilities. The computation of $\Pi(X_i|$ Par(X_i)) is immediately obtained from Σ_{Xi}, as it is given by the following way: Let $\{\Sigma_{X1}, ..., \Sigma_{Xn}\}$ be the result of Step 6 of the previous algorithm. Let X_i be a variable and Par (X)=$\{Y_1, ..., Y_m\}$ be the set of its parents. Let x be an instance of X_i and $P_X=y_1 \wedge...\wedge y_m$ be an instance of Par (X_i). Let:
    $\Pi(x|P_X) = 1-a_i$ if $(\neg x \vee \neg P_X, a_i) \in \Sigma_{Xi}$
        = 1 otherwise.
Then we have:

Proposition 13: The possibility distribution associated to Σ and the possibility distribution obtained from the graph using the minimum operator are equal, namely: $\pi_\Sigma = \pi_m$.

Example 10 (continued) :

Let us show how compute $\pi(\neg a|\neg b\neg cd)$ and $\pi(\neg a|\neg b\neg c\neg d)$, the others are obtained in the same way. Recall that

$\Sigma_A = \{(a\vee b\vee c\vee d\ 0.7), (a\vee b\vee c\vee\neg d\ 0.7), (a\vee b\vee\neg c\vee d\ 0.7), (a\vee b\vee\neg c\vee\neg d\ 0.7), (\neg a\vee b\vee c\vee\neg d\ 0.7), (\neg a\vee\neg b\vee c\vee\neg d\ 0.7), (a\vee b\vee c\vee d\ 0.9), (a\vee\neg b\vee c\vee d\ 0.9)\}$.

Then by definition:

$$\pi(\neg a|\neg b\neg cd) = .3$$
$$\pi(\neg a|\neg b\neg c\neg d) = .1$$

The conditional possibility distributions are given by the following tables :

$\pi(A \mid BCD)$:

A | BCD	$\neg b\neg cd$	$\neg bc\neg d$	$\neg bcd$	$\neg b\neg cd$
a	1	1	1	0.3
$\neg a$	0.3	0.3	0.3	1

A | BCD	$b\neg cd$	$\neg b\neg c\neg d$	$b\neg c\neg d$	otherwise
a	0.3	1	1	1
$\neg a$	1	0.1	0.1	1

$\pi(B \mid CE)$:

B | CE	$\neg c\neg e$	$\neg ce$	$\neg c\neg e$	$c\neg e$
b	1	1	0.8	0.8
$\neg b$	0.2	0.2	1	1

$\pi(D \mid \neg F)$:

D | F	f	$\neg f$
d	1	0.5
$\neg d$	1	1

$\pi(C)$:

c	1
$\neg c$	1

$\pi(E)$:

e	1
$\neg e$	1

$\pi(F)$:

f	1
$\neg f$	1

Let us notice that it is possible to express ignorance, i.e., it is not necessary to give the a priori possibilities on X if they are unknown.

Proposition 14: Let $\Pi(X|Par(X))$ be the conditional possibility distributions over the variables X in the DAG associated to Σ. Let π_m be the joint possibility distribution obtained using the chain rule with minimum from a causal network. Then: $\pi_m(a|Pa) = \Pi(a|P_a)$.

The previous proposition shows that the constructed DAG from the possibilistic logic base guarantees the recovering of initial values, when using the min-based chain rule, and hence it satisfies the independence relations in the sense of Definition 2 encoded by the structure of the graph.

6. Conclusion

This paper has established the links between possibilistic logic and directed possibilistic graphs. We have shown that directed possibilistic graphs can be encoded into possibilistic logic, for the two possible definitions of conditioning based on the minimum and the product operator. When it is based on product, it also provides a

mean for turning a Bayesian-net equipped with a kappa-function into a possibilistic logic base (using the transformation possibility measures-kappa functions recalled at the end of Section 2.2.).

The inverse passage from a possibilistic logic base to a network has also been provided. This allows the expert to express his knowledge using "causality" relations between variables, and then the possibilistic logic machinery can be applied after the computation of the corresponding possibilistic logic base.

A future work would be the study of the complexity of the conversion of a directed causal network based on the product into possibilistic logic (with the min, the conversion is linear) and the comparison of the cost of the inference using the network directly with the one using the corresponding possibilistic knowledge base, which may be also used for explanation purposes.

References

[1] S. Benferhat, D. Dubois, L. Garcia and H. Prade (1998) Directed possibilistic graphs and possibilistic logic. Proceedings of IPMU-98, pp. 1470-1477.

[2] S. Benferhat, D. Dubois, L. Garcia and H. Prade (1999) *Possibilistic logic bases and possibilistic graphs.* Technical Report, IRIT, Univ. P. Sabatier, Toulouse.

[3] S. Benferhat, D. Dubois and H. Prade (1997). Syntactic combination of uncertain information: a possibilistic approach. *Proc. of Qualit. and Quantit. Practical Reasoning (ECSQARU-FAPR'97).* LNAI 1244,pp. 30-42.

[4] L. M. de Campos et J. F. Huete (1994). *Independence in possibility theory.* Department of Artificial Intelligence. University of Granada, Spain.

[5] A. P. Dawid (1979). Conditional independence in statistical theory. *J. Royal Statistical Society A*, 41. 1-31.

[6] D. Dubois, L. Fariñas del Cerro, A. Herzig and H. Prade (1994). An ordinal view of independence with application to plausible reasoning. *Proc. of UAI-94* , 195-203.

[7] D. Dubois, J. Lang and H. Prade (1994). Possibilistic logic. In: *Handbook of Logic in Artificial Intelligence and Logic Programming*, Vol. 3 (D.M. Gabbay et al., eds.). Oxford University Press. 439-513.

[8] Dubois D., Prade H. (1998) Possibility theory: qaulitative and quantitative aspects. In *Handbook of Defeasible Reasoning and Uncertainty Management Systems.* Vol. 1, pp. 169-226, Kluwer Academic Press.

[10] P. Fonck (1993). *Réseaux d'inférence pour le traitement possibiliste.* Thèse de doctorat. Univ. de Liège.

[11] J. Gebhardt and R. Kruse (1997). Background and perspectives of possibilistic graphical models. *Proc. of Qualit. and Quantit. Practical Reasoning (ECSQARU-FAPR'97).* LNAI 1244. Springer. 108-121.

[12] E. Hisdal (1978). Conditional possibilities, independence and noninteraction. *Fuzzy Sets and Systems*, 1. 283-297.

[13] F. V. Jensen (1996). *An Introduction to Bayesian Networks.* University College London: UCL Press.

[14] J. Pearl (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.* Morgan Kaufmann.

[15] W. Spohn (1988) Ordinal conditional functions: A dynamic theory of epistemic states. In: *Causation in Decision, Belief Change, and Statistics,* Vol. 2 (W.L. Harper, B. Skyrms, eds.), D. Reidel, Dordrecht, 105-134.

[16] L.A. Zadeh (1978). Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 1. 3-28.