Converting Possibilistic Networks by Using Uncertain Gates

Guillaume Petiot( )

CERES, Catholic Institute of Toulouse, 31 rue de la Fonderie,
31068 Toulouse, France

guillaume.petiot@ict-toulouse.fr

<http://www.ict-toulouse.fr>

Abstract. The purpose of this paper is to define a general frame to convert the Conditional Possibility Tables (CPT) of an existing possibilistic network into uncertain gates. In possibilistic networks, CPT parameters must be elicited by an expert but when the number of parents of a variable grows, the number of parameters to elicit grows exponentially. This problem generates difficulties for experts to elicit all parameters because it is time-consuming. One solution consists in using uncertain gates to compute automatically CPTs. This is useful in knowledge engineering. When possibilistic networks already exist, it can be interesting to transform them by using uncertain gates because we can highlight the combination behaviour of the variables. To illustrate our approach, we will present at first a simple example of the estimation for 3 test CPTs with behaviours MIN, MAX and weighted average. Then, we will perform a more significant experimentation which will consist in converting a set of Bayesian networks into possibilistic networks to perform the estimation of CPTs by uncertain gates.

Keywords: Possibilistic networks · Possibility theory · Uncertain logical gates · Estimation

1 Introduction

Knowledge engineering often needs to describe how information is combined. Experts' knowledge can be represented by a Directional Acyclic Graph (DAG). Unfortunately, this kind of knowledge is often imprecise and uncertain. The use of possibility theory [21] allows us to take into account these imperfections. Possibilistic networks [1, 2], as Bayesian networks [17, 20] in probability theory, can be used to propagate new information, called evidence, in a DAG. The main disadvantage of possibilistic networks is the need to elicit all parameters of Conditional Possibility Tables. When the number of parents of a variable grows, the number of parameters to elicit in the CPT grows exponentially. For example, if a variable of 2 states has 10 parents with 2 states, then we have $2^{11} = 2048$ parameters to elicit. Moreover, when experts define the parameters

of a CPT in possibilistic networks, they often define an implicit behaviour. The most common behaviours are those of Boolean logic, such as AND, OR, XOR, etc.

Noisy gates in probability theory [4] deal with this problem of parameters. They use a function to compute automatically the CPT. As a result, the number of parameters is reduced. Another advantage is the modelling of noise and incomplete knowledge. Indeed, it is often difficult to have all the knowledge during the task of modelling because of the complexity of the problem. The use of a variable which represents unknown knowledge leads to another kind of model. In possibility theory, authors Dubois et al. [5] proposed to use uncertain logical gates which present the same advantages as noisy gates. The behaviour of uncertain gates connectors depends on the choice of the function. The function can have the following behaviours: indulgent, compromise or severe.

Nowadays, there are several applications of possibilistic networks that exist where experts have elicited all CPT parameters. We propose in this paper a solution to convert an existing possibilistic network into a possibilistic network with uncertain gates. It is very interesting to see which connector corresponds to each CPT. This can provide a better understanding of information processing, which is not highlighted when an expert elicits CPT parameters. This is also useful in knowledge engineering.

This paper is structured as follows. In the first part, we will present possibility theory and uncertain gates. Then we will focus our interest on the estimation of uncertain gates and we will propose the examples of the estimation of uncertain gates for test CPTs and existing applications using Bayesian networks. So we will have to convert Bayesian networks into possibilistic networks and then we will perform the estimation of the CPTs. We will present the results of this experiment in the last part of this paper.

2 Uncertain Gates

Uncertain gates are an analogy of noisy gates in possibility theory [21]. Possibility theory proposes to model imprecise knowledge by using a possibility distribution. If Ω is the referential, then the possibility distribution π is defined from Ω to $[0, 1]$ with $\max_{v \in \Omega} \pi(v) = 1$. Dubois et al. [6] define the possibility measure Π and the necessity measure N from $P(\Omega)$ to $[0, 1]$:

$$\forall A \in P(\Omega), \Pi(A) = \sup_{x \in A} \pi(x) \text{ and } N(A) = 1 - \Pi(\neg A) = \inf_{x \notin A} 1 - \pi(x). \quad (1)$$

The possibility theory is not additive but maxitive:

$$\forall A, B \in P(\Omega), \Pi(A \cup B) = \max(\Pi(A), \Pi(B)). \quad (2)$$

We can compute the possibility of a variable A by using the conditioning proposed by D. Dubois and H. Prade [6]:

$$\forall A, B \in P(\Omega), \Pi(A|B) = \begin{cases} \Pi(A, B) & \text{if } \Pi(A, B) < \Pi(B), \\ 1 & \text{if } \Pi(A, B) = \Pi(B). \end{cases} \quad (3)$$

Possibilistic networks [1, 2] can be defined as follows for a DAG $G = (V, E)$ where V is the set of Variables and E the set of edges between the variables:

$$\Pi(V) = \bigotimes_{X \in V} \Pi(X/Pa(X)). \quad (4)$$

Pa stands for the parents of the variable X . There are two kinds of possibilistic networks: qualitative possibilistic networks (also called min-based possibilistic networks) where $\bigotimes$ is the function min, and quantitative possibilistic networks (also called product-based possibilistic networks) where $\bigotimes$ is the product. In this research, we consider the comparison of possibilistic values instead of the use of an intensity scale in $[0, 1]$ so we will use a min-based possibilistic network.

We can propose as in noisy gates to use the Independence of Causal Influence [4, 9, 25] to define a possibilistic model with the ICI. In fact, as in probability theory, there is a set of causal variables $X_1, \dots, X_n$ which influence the result of an effect variable Y . We insert between each X_i s and Y a variable Z_i to take into account the noise, i.e. uncertainty. We can also introduce another variable Z_l which represents the unknown knowledge [5]. The following figure presents the possibilistic model with ICI (Fig. 1):

Fig. 1. Possibilistic model with ICI.

We propose to compute $\pi(Y|X_1, \dots, X_n)$ by marginalizing the variables Z_i s:

$$\pi(y|x_1, \dots, x_n) = \bigoplus_{z_1, \dots, z_n} \pi(y|z_1, \dots, z_n) \bigotimes \bigotimes_{i=1}^n \pi(z_i|x_i) \quad (5)$$

The $\bigotimes$ is the minimum and the $\bigoplus$ is the maximum in possibility theory. If we have a deterministic model where the value of Y can be computed from the values of Z_i s by using a function f then:

$$\text{where } \pi(y|z_1, \dots, z_n) = \begin{cases} 1 & \text{if } y = f(z_1, \dots, z_n) \\ 0 & \text{else} \end{cases} \quad (6)$$

As a result, we obtain the following equation:

$$\pi(y|x_1, \dots, x_n) = \bigoplus_{z_1, \dots, z_n: y=f(z_1, \dots, z_n)} \bigotimes_{i=1}^n \pi(z_i|x_i) \quad (7)$$

If we add a leakage variable Z_l in the previous equation, we obtain:

$$\pi(y|x_1, \dots, x_n) = \bigoplus_{z_1, \dots, z_n, z_l: y=f(z_1, \dots, z_n, z_l)} \bigotimes_{i=1}^n \pi(z_i|x_i) \otimes \pi(z_l) \quad (8)$$

The CPT is computed by using this formula. The function f can be AND, OR or their generalizations uncertain MIN and uncertain MAX as proposed by Dubois et al. [5]. These authors developed an optimization for the computation of uncertain MIN and MAX connectors. We can also use a weighted average function or the operator OWA [22], as described in our previous study [13, 14]. For this experimentation, we will use the connectors uncertain MIN (UMIN), uncertain MAX (UMAX) and uncertain weighted average (UWAVG). The function f must have the same domain as the variable Y . We can see that the connectors uncertain MIN and uncertain MAX satisfy this property. The weighted average of the intensity level of the variables X_i can return a value outside the domain of Y . So in this case we have to combine a function noted g with a scaling function f_s which returns a value in the domain of Y . We can write $f = f_s \circ g$ where function g is $g(z_1, \dots, z_n) = \omega_1 z_1 + \dots + \omega_n z_n$. If the state of the variable Y defines an ordered scale $E = \{\vartheta_0 < \vartheta_1 < \dots < \vartheta_m\}$, the function f_s can be for example:

$$f_s(x) = \begin{cases} \vartheta_0 & \text{if } x \leq \theta_0 \\ \vartheta_1 & \text{if } \theta_0 < x \leq \theta_1 \\ \vdots & \vdots \\ \vartheta_m & \text{if } \theta_{m-1} < x \end{cases} \quad (9)$$

The parameters θ_i allow us to adjust the behaviour of f_s . The function g has n parameters which are the weights w_i of the linear combination and n arguments. If all weights are equal to $\frac{1}{n}$, then we calculate the average of the intensities. If $\forall_{i \in [1, n]} \omega_i = 1$, then we make the sum of the intensities. To compute the Eq. 9, we have to define $\pi(Z_i|X_i)$, $\pi(Z_l)$, and the function f . If the states of a variable are ordered, we can assign an intensity level to each state as done by Dubois et al. [5]. In our experimentation with test CPTs we propose to use 3 states for the variables such as low, medium and high. The intensity levels are 0 for low, 1 for medium and 2 for high. The table of $\pi(Z_i|X_i)$ is the following (Table 1):

Table 1. Possibility table for 3 ordered states.

$\pi(Z_i X_i)$	$x_i = 2$	$x_i = 1$	$x_i = 0$
$z_i = 2$	1	$s_i^{2,1}$	$s_i^{2,0}$
$z_i = 1$	$\kappa_i^{1,2}$	1	$s_i^{1,0}$
$z_i = 0$	$\kappa_i^{0,2}$	$\kappa_i^{0,1}$	1

In the previous table, κ represents the possibility that an inhibitor exists if the cause is met and s_i the possibility that a substitute exists when the cause is not met. If a cause of weak intensity cannot produce a strong effect, then all

$s_i = 0$. So there are 6 parameters at the most per variable and 2 parameters for $\pi(Z_l)$. If all variables have the same intensity level and are graded variables with a normal state of intensity 0 (the first state is the normal state), there is no problem. In this case, the higher the intensity is, the more abnormal the situation becomes. But there are two other cases to discuss. The first case concerns the variables which are not graded variables. For example, if we have a variable where the domain is (*decrease, normal, accelerate*), we can see that the normal speed is the neutral state i.e. with the intensity level 1. The second case concerns the incompatibilities of the intensity levels of the causal variables X_i s and their parents. Indeed, if some causal variables do not share the same domain with the effect variable, we can imagine that the computation of the function f can give the results that are out of range. This remark leads to a constraint on Z_i s, which is that Z_i s must have the same domain as Y .

3 Estimation

The estimation of uncertain gates leads to two problems discussed in [24]. The first one is the simple case where we would like to perform an estimation of uncertain gates from an existing CPT. The second case concerns the estimation of the connector from the data without any CPT. Indeed, in possibilistic networks, the number of parameters to elicit grows exponentially with the number of parents of a variable, leading to problems of knowledge engineering. So if data is available, we propose to perform the estimation of the CPT instead of eliciting all parameters. Unfortunately, it is often difficult to have enough data to estimate all parameters.

In this paper, we will focus our interest on the first case. If a CPT already exists, it can be useful to know which uncertain gate corresponds to the CPT. The question can be which connector matches the target CPT among all existing connectors. To discover this, we have to compare the CPT generated by uncertain gates and the existing CPT. This leads to the problem of estimating the parameters of uncertain gates. We look for the closest CPT to a reference CPT. To do this, we must use a distance [8]. For our first experimentation we will use the Euclidean distance. The problems discussed in [24] for this distance are also present in our research. We propose to analyze them and to compare several distances in future works. With respect to this study, we will compare 4 estimation methods by using the distance of the CPTs as a cost function in order to improve the estimation: hill-climbing (HC), simulated annealing (SA), tabu search (TS), and genetic algorithm (GA). We propose to choose the connector with the smallest distance $\hat{\theta}$ to the target CPT.

We consider the example of hill-climbing and the estimation of the parameters of the uncertain MIN. We must at first initialize all parameters of the uncertain MIN to a random value in $[0, 1]$. Then we generate the CPT by using the connector and we calculate the initial distance between the generated CPT and the target CPT. Next, we evaluate all neighbours of the parameters by adding and subtracting a step (Function *GenNeighbour* in the following algorithm). We consider only the parameters with the smallest distance. These parameters become

our temporary solution. We reiterate this process until the distance is lower than a constant, or a maximum number of iterations is reached. The algorithm for the estimation of the parameters of the uncertain MIN with hill-climbing is the following:

Algorithm 1: Hill climbing.

Input : Y a variable; $X_1, \dots, X_n$ the n parents of Y ; $\pi(Y|X_1, \dots, X_n)$ the initial CPT; max the maximum number of iterations; ϵ the accepted distance; $step$ a constant.

Output: The result is $\pi(Z_i|X_i)$.

```

1 begin
2   iteration  $\leftarrow 0$ ; error  $\leftarrow +\infty$ ; current  $\leftarrow +\infty$ ; Initialize( $\pi(Z_i|X_i)$ );
3   while iteration < max and error >  $\epsilon$  do
4      $\pi'(Z_i|X_i) \leftarrow \pi(Z_i|X_i)$ ;
5     current  $\leftarrow$  error;
6     forall  $\pi^*(Z_i|X_i) \in GenNeighbour(\pi(Z_i|X_i), step)$  do
7        $\pi^*(Y|X_1, \dots, X_n) \leftarrow UMIN(Y, \pi^*(Z_i|X_i))$ ;
8       if  $E(\pi^*(Y|X_1, \dots, X_n), \pi(Y|X_1, \dots, X_n)) < current$  then
9         current  $\leftarrow E(\pi^*(Y|X_1, \dots, X_n), \pi(Y|X_1, \dots, X_n))$ ;
10         $\pi'(Z_i|X_i) \leftarrow \pi^*(Z_i|X_i)$ ;
11    if current < error then
12      error  $\leftarrow$  current;
13       $\pi(Z_i|X_i) \leftarrow \pi'(Z_i|X_i)$ ;
14    iteration  $\leftarrow$  iteration + 1;
```

The simulated annealing algorithm was proposed [18] to describe a thermodynamic system. This is a probabilistic approach which leads to an approximation of a global optimum. We have proposed the following algorithm for the estimation of the parameters of the uncertain MIN:

Algorithm 2: Simulated annealing.

Input : Y a variable; $X_1, \dots, X_n$ the n parents of Y ; $\pi(Y|X_1, \dots, X_n)$ the initial CPT; max the maximum number of iterations; ϵ the accepted distance; $step$ a constant; K_{max} is the maximum of steps; T is the temperature.

Output: The result is $\pi(Z_i|X_i)$.

```

1 begin
2   iteration  $\leftarrow 0$ ; Initialize( $\pi(Z_i|X_i)$ ); SP  $\leftarrow +\infty$ ;
3    $\pi^p(Y|X_1, \dots, X_n) \leftarrow UMIN(Y, \pi(Z_i|X_i))$ ;
4   while iteration < max and SP >  $\epsilon$  do
5     k  $\leftarrow 0$ ;
6     while k <  $k_{max}$  do
7        $\pi^*(Z_i|X_i) \leftarrow GenNeighbour(\pi(Z_i|X_i), step)$ 
8        $\pi^*(Y|X_1, \dots, X_n) \leftarrow UMIN(Y, \pi^*(Z_i|X_i))$ ;
9       SP  $\leftarrow E(\pi^p(Y|X_1, \dots, X_n), \pi(Y|X_1, \dots, X_n))$ ;
10      SN  $\leftarrow E(\pi^*(Y|X_1, \dots, X_n), \pi(Y|X_1, \dots, X_n))$ ;
11      delta  $\leftarrow$  SN - SP;
12      if delta < 0 then
13         $\pi(Z_i|X_i) \leftarrow \pi^*(Z_i|X_i)$ ;
14         $\pi^p(Y|X_1, \dots, X_n) \leftarrow \pi^*(Y|X_1, \dots, X_n)$ ;
15        r  $\leftarrow RandomValue()$ ;
16        if  $r < e^{-\frac{delta}{T}}$  then
17           $\pi(Z_i|X_i) \leftarrow \pi^*(Z_i|X_i)$ ;
18           $\pi^p(Y|X_1, \dots, X_n) \leftarrow \pi^*(Y|X_1, \dots, X_n)$ ;
19        k  $\leftarrow$  k + 1;
20    T  $\leftarrow \alpha \times T$ ;
21    iteration  $\leftarrow$  iteration + 1;
```

We also propose to use a genetic algorithm to perform the estimation. The chromosomes gather all the parameters of the connector. The population of

chromosomes will evolve by favouring the best individuals. A fitness function, which measures the distance between a CPT generated by one chromosome and the target CPT, is used to compare the individuals. The best chromosome is the one with the smallest fitness. The genetic algorithm performs several processing operations in a loop: the selection of the parents, the crossover, and the mutation. We can resume this with the following algorithm:

Algorithm 3: Genetic algorithm.

Input : Y a variable; $X_1, \dots, X_n$ the n parents of Y ; $\pi(Y|X_1, \dots, X_n)$ the initial CPT; max the maximum number of iterations; ϵ the accepted distance; $step$ a constant; N is the number of chromosomes;
Output: The result is $\pi(Z_i|X_i)$.

```

1 begin
2    $generation \leftarrow 0$ ;  $bestFitness \leftarrow +\infty$ ;  $InitializePopulation()$ ;
3   while  $generation < max$  and  $bestFitness > \epsilon$  do
4      $ComputeFitness()$ ;
5      $TournamentSelection()$ ;
6      $Crossover()$ ;
7      $Mutation()$ ;
8      $generation \leftarrow generation + 1$ ;

```

The tabu search was proposed by F. W. Glover [10–12]. This algorithm improves local search algorithms by using a list of previous moves to avoid processing a situation already explored. The Tabu search algorithm is the following:

Algorithm 4: Tabu search.

Input : Y a variable; $X_1, \dots, X_n$ the n parents of Y ; $\pi(Y|X_1, \dots, X_n)$ the initial CPT; max the maximum number of iterations; ϵ the accepted distance; $step$ a constant; NC is the number of candidates; T is the temperature.
Output: The result is $\pi(Z_i|X_i)$.

```

1 begin
2    $iteration \leftarrow 0$ ;  $Initialize(\pi(Z_i|X_i))$ ;  $\pi^p(Y|X_1, \dots, X_n) \leftarrow UMIN(Y, \pi(Z_i|X_i))$ ;
3    $bestScore \leftarrow +\infty$ ;
4   while  $iteration < max$  and  $bestScore > \epsilon$  do
5      $k \leftarrow 0$ ;
6     while  $k < NC$  do
7        $\pi^*(Z_i|X_i) \leftarrow GenNeighbour(\pi(Z_i|X_i), step)$ 
8        $\pi^*(Y|X_1, \dots, X_n) \leftarrow UMIN(Y, \pi^*(Z_i|X_i))$ ;
9        $score \leftarrow E(\pi^*(Y|X_1, \dots, X_n), \pi(Y|X_1, \dots, X_n))$ ;
10      if  $\pi^*(Z_i|X_i) \notin Candidate$  then
11         $Candidate \leftarrow Candidate \cup \pi^*(Z_i|X_i)$ ;
12         $F(k) \leftarrow score$ ;
13         $k \leftarrow k + 1$ ;
14       $S \leftarrow IndicesOfAscendingSort(F)$ 
15      if  $(Candidate[S[0]] \notin ListTabu)$  or  $((Candidate[S[0]] \in ListTabu)$  and
16         $(F(S[0]) < bestScore))$  then
17         $ListTabu \leftarrow ListTabu \cup Candidate[S[0]]$ ;
18         $\pi(Z_i|X_i) \leftarrow Candidate[S[0]]$ ;
19         $currentScore \leftarrow F(S[0])$ ;
20      else
21         $b \leftarrow 0$ ;
22        while  $((b < NC)$  and  $(Candidate[S[0]] \in ListTabu))$  do
23           $b \leftarrow b + 1$ ;
24        if  $b < NC$  then
25           $ListTabu \leftarrow ListTabu \cup Candidate[S[b]]$ ;
26           $\pi(Z_i|X_i) \leftarrow Candidate[S[b]]$ ;
27           $currentScore \leftarrow F(S[b])$ ;
28        if  $currentScore < bestScore$  then
29           $bestScore \leftarrow currentScore$ ;
30       $iteration \leftarrow iteration + 1$ ;

```

4 Experimentation

We now propose to test this approach on simulated CPTs for connectors MIN, MAX and weighted average. Next, we will compare the estimation of the CPTs. For this experimentation, we will use only two causal variables X_1 and X_2 and one effect variable Y . All the variables have states: low, medium and high with intensity levels 0, 1 and 2. We provide below the parameters used to compute the simulated CPTs (Table 2):

Table 2. Possibility table of $\pi(Z_i|X_i)$.

$\pi(Z_i X_i)$	$x_i = 2$	$x_i = 1$	$x_i = 0$
$z_i = 2$	1	0.3	0
$z_i = 1$	0.3	1	0
$z_i = 0$	0.2	0.3	1

We used the same parameters for X_1 and X_2 . We defined $\pi(Z_i)$ with the values $\pi(Z_L = 0) = 1$ and $\pi(Z_L = 1) = \pi(Z_L = 2) = 0.1$. For the weighted average, we will simulate a mean behaviour by defining the weights equal to 0.5 for all variables but we will not take into account the leakage variable. We present the simulated CPTs obtained by using the above parameters in the following tables (Table 3):

Table 3. Simulated CPTs.

		X_1			Low			Medium			High		
		X_2			Low	Medium	High	Low	Medium	High	Low	Medium	High
Y	Low				1.0	1.0	1.0	1.0	0.3	0.3	1.0	0.3	0.2
	Medium				0.1	0.1	0.1	0.1	1.0	1.0	0.1	1.0	0.3
	High				0.1	0.1	0.1	0.1	0.3	0.3	0.1	0.3	1.0

(a) MIN CPT.

		X_1			Low			Medium			High		
		X_2			Low	Medium	High	Low	Medium	High	Low	Medium	High
Y	Low				1.0	0.3	0.2	0.3	0.3	0.2	0.2	0.2	0.2
	Medium				0.1	1.0	0.3	1.0	1.0	0.3	0.3	0.3	0.3
	High				0.1	0.3	1.0	0.3	0.3	1.0	1.0	1.0	1.0

(b) MAX CPT.

		X_1			Low			Medium			High		
		X_2			Low	Medium	High	Low	Medium	High	Low	Medium	High
Y	Low				1.0	1.0	0.3	1.0	0.3	0.3	0.3	0.3	0.2
	Medium				0.0	0.3	1.0	0.3	1.0	0.3	1.0	0.3	0.3
	High				0.0	0.0	0.0	0.0	0.3	1.0	0.0	1.0	1.0

(c) WAVG CPT.

We have performed the estimation with several steps 0.1, 0.01 and 0.001 and a limited number of iteration of 40000. The result of the estimation is the following (Table 4):

Table 4. Result of the estimation.

Connectors	Uncertain MIN				Uncertain MAX				Uncertain WAVG			
	HC	SA	TS	GA	HC	SA	TS	GA	HC	SA	TS	GA
MIN CPT	0.31	0.0	0.0	0.0	2.40	2.27	2.27	2.27	1.14	0.86	0.86	0.86
MAX CPT	2.22	2.05	2.11	2.17	0.62	0.0	0.0	0.22	1.09	0.94	0.86	0.86
WAVG CPT	1.84	1.72	1.72	1.73	2.08	1.96	1.96	1.98	0.10	0.0	0.22	0.0

(a) *step* = 0.1.

Connectors	Uncertain MIN				Uncertain MAX				Uncertain WAVG			
	HC	SA	TS	GA	HC	SA	TS	GA	HC	SA	TS	GA
MIN CPT	0.27	0.0	0.24	0.07	2.39	2.27	2.43	2.27	1.93	1.21	1.04	1.28
MAX CPT	2.18	2.05	2.08	2.08	0.26	0.0	0.0	0.0	1.63	2.21	1.29	0.86
WAVG CPT	1.89	1.72	1.72	1.92	2.22	2.01	1.96	1.96	0.56	1.24	0.08	0.22

(b) *step* = 0.01.

Connectors	Uncertain MIN				Uncertain MAX				Uncertain WAVG			
	HC	SA	TS	GA	HC	SA	TS	GA	HC	SA	TS	GA
MIN CPT	0.27	0.0	2.15	0.07	2.39	2.69	2.29	2.27	1.93	1.04	1.04	1.81
MAX CPT	2.11	2.08	2.25	2.10	0.20	0.58	0.0	0.22	1.86	1.80	1.63	0.95
WAVG CPT	1.89	1.73	2.57	1.73	2.22	1.97	1.96	1.97	0.58	0.22	0.002	0.08

(c) *step* = 0.001.

We can see in the above tables that all simulated CPTs are associated with the expected connector. We can see in bold in each line the smallest distance which corresponds to the best result for the estimation. By applying the decision rule, we select for each line the connector and the estimation algorithm which has produced the best result. For example, with step 0.001 we can see that the simulated MIN CPT was associated with the connector uncertain MIN and that the best result was provided by the simulated annealing algorithm with the final distance of 0. This means that the parameters are exactly estimated. The simulated MAX CPT was associated to the connector uncertain MAX and the best result was provided by tabu search. And finally, the simulated WAVG CPT was associated to uncertain WAVG connector as expected and the best result was provided by tabu search with a distance of 2×10^{-3} . If we compare the results, we can see that the best results are obtained with step 0.1. So we provide below the comparison of the estimation algorithms which lead to the best results for this step (Fig. 2):

Fig. 2. Estimation of the simulated CPT with *step* = 0.1 for the first 10,000 iterations.

We can see in the above results that all algorithms did not converge to 0. The best result seems to be obtained by simulated annealing or tabu search but for *step*=0.001 and for the simulated MIN CPT the tabu algorithm didn't find the expected parameters. Nevertheless, we have tried several parameters for tabu search and we have obtained good results.

We have performed another experimentation by using existing Bayesian networks: Asia [16], Earthquake and Cancer [15], Survey [19], Sachs [23], Andes [3] because of the small number of data set available for possibilistic networks. We have transformed the conditional probability tables into conditional possibility tables to obtain possibilistic networks. The structure of the networks is the same in Bayesian networks and possibilistic networks. Then we have applied our approach to replace all CPTs by appropriate connectors. The problem of converting the probability table has already been discussed by Dubois et al. [7]. The temptation can be to perform the normalization of the probability but it is not sufficient to ensure that the possibility property $\Pi \geq P$ is satisfied. For example, if we have $p_1 > p_2 > \dots > p_n$, we perform $p_i = \frac{p_i}{p_1}$ for all i . The correct solution proposed by Dubois et al. [7] consists in computing the possibilities as follows:

$$\pi_i = \begin{cases} 1 & \text{if } i = 1, \\ \sum_{j=i}^n p_j & \text{if } \pi_{i-1} > \pi_i, \\ \pi_{i-1} & \text{otherwise.} \end{cases} \quad (10)$$

We present below the result of this experimentation. At first, we have performed the estimation for small networks. We can see in the first column the name of the data set, then the name of the node and the number of the parents of the node. In the fourth column, we present the size of the CPT, and in the following columns, the minimal distance for all connectors (Table 5).

Table 5. Results for small size networks (*step* = 0.01).

Name	Table	Parents	Size	MIN	MAX	WAVG
Asia	Dysp	2	8	1.238	0.0707	0.0709
	Either	2	8	1.581	0.011	0.012
Earthquake	Alarm	2	8	0.961	0.71	0.204
Cancer	Cancer	2	8	0.97	1.6856	0.07
Survey	E	2	12	1.56	0.71	0.71
	T	2	12	0.70	1.06	0.74

This first experiment allows us to distinguish two cases. The first one concerns the distance close to 0. In this case, the computation of the connector matches the target CPT. The connector can replace the CPT. The second case concerns the distance not close enough to 0, such as the tables alarm, E or T. We cannot replace the CPT by a connector, nevertheless the information provided is meaningful because we can deduce the behaviour of the CPT between severe and indulgent. We have also performed another experiment with Andes, which is a big network, and we obtain the following result for the first 30 nodes (Table 6):

Table 6. The 30 first nodes of Andes (*step* = 0.01).

Name	Table	Parents	Size	MIN	MAX	WAVG
Andes	RApp1	2	8	0.01	1.58	0.01
	SNode_8	2	8	1.52	0.007	0.008
	SNode_20	2	8	0.005	1.33	0.006
	SNode_21	2	8	0.006	1.33	0.003
	SNode_24	2	8	0.009	1.45	0.008
	SNode_25	2	8	0.008	1.45	0.001
	SNode_26	2	8	0.006	1.33	0.007
	SNode_47	2	8	0.003	1.33	0.007
	RApp3	3	16	0.027	2.598	0.013
	RApp4	2	8	0.015	1.581	0.01
	SNode_27	2	8	1.523	0.006	0.004
	GOAL_48	2	8	0.006	1.33	0.003
	GOAL_49	4	32	0.008	3.13	0.01
	GOAL_50	2	8	0.006	1.33	0.009
	SNode_51	3	16	0.009	2.36	0.009
	SNode_52	2	8	0.004	1.33	0.008
	GOAL_53	3	16	0.01	2.14	0.008
	SNode_28	3	16	0.0104	2.14	0.0105
	SNode_29	2	8	0.006	1.33	0.003
	SNode_31	2	8	0.009	1.45	0.008
	SNode_33	2	8	0.009	1.45	0.006
	SNode_34	2	8	0.004	1.33	0.007
	GOAL_56	3	16	0.013	2.14	0.009
	GOAL_57	2	8	0.005	1.33	0.003
	SNode_59	3	16	0.006	2.14	0.008
	SNode_60	2	8	0.003	1.33	0.006
	GOAL_61	3	16	0.009	2.14	0.007
	GOAL_62	2	8	0.005	1.33	0.006
	GOAL_63	3	16	0.01	2.14	0.008
	SNode_64	2	8	0.007	1.33	0.006

Having computed all nodes, we obtain the following results: for the uncertain MIN connector, 94% of the final distance is less than 0.1; for the uncertain MAX connector, only 6 results are below 0.1; for the uncertain WAVG connector most of the results have a distance less than 0.1 because it can replace the other connectors. Nevertheless, its computation time is high. The best results for the uncertain MIN connector were obtained by using the hill-climbing algorithm and the genetic algorithm. Nevertheless, the use of simulated annealing and

tabu search can sometimes provide good results too. As for the uncertain MAX and uncertain WAVG connectors, all algorithms gave the same number of good results. The use of concurrent algorithms for the estimation has improved the final result.

5 Conclusion

In this paper, we have proposed a solution to convert the CPTs of an existing possibilistic network into uncertain gates. Uncertain gates provide a set of connectors with behaviours from severe to indulgent through the family of compromise (median, average, weighted average,...). Uncertain gates provide meaningful knowledge about how information is combined to compute the state of a variable. To convert the CPTs, we have used and compared optimization algorithms such as hill-climbing, simulated annealing, tabu search and the genetic algorithm. The goal was to find the closest CPT generated by using the connectors. The best result is provided by the connector which matches the best the initial CPT. To validate this approach, we have generated test CPTs with the behaviours MIN, MAX and weighted average and performed the estimation. The results correspond to our expectations. We have also proposed to test our solution on real datasets. To do this, we have converted Bayesian networks into possibilistic networks and performed the estimation of the connectors to select the most appropriate one for each table. For our future works we would like to improve the computation of the connectors by proposing a parallel algorithm. We would like to better analyze the problem of variables with different intensity scales. Also, it can be interesting to evaluate the effects on the decision of converting a CPT into uncertain gates.

References

1. Benferhat, S., Dubois, D., Garcia, L., Prade, H.: Possibilistic logic bases and possibilistic graphs. In: *Proceedings of the Fifteenth Conference on Uncertainty in Artificial Intelligence*, pp. 57–64, Morgan Kaufmann, Stockholm (1999)
2. Borgelt, C., Gebhardt, J., Kruse, R.: Possibilistic graphical models. In: Della Riccia, G., Kruse, R., Lenz, H.-J. (eds.) *Computational Intelligence in Data Mining. ICMS*, vol. 408, pp. 51–67. Springer, Vienna (2000). https://doi.org/10.1007/978-3-7091-2588-5_3
3. Conati, C., Gertner, A.S., VanLehn, K., Druzdzel, M.J.: On-line student modeling for coached problem solving using Bayesian networks. In: Jameson, A., Paris, C., Tasso, C. (eds.) *User Modeling. ICMS*, vol. 383, pp. 231–242. Springer, Vienna (1997). https://doi.org/10.1007/978-3-7091-2670-7_24
4. Diez, F., Drudzel, M.: Canonical probabilistic models for knowledge engineering. Technical Report, CISIAD-06-01 (2007)
5. Dubois, D., Fusco, G., Prade, H., Tettamanzi, A.: Uncertain logical gates in possibilistic networks. an application to human geography. In: Beierle, C., Dekhtyar, A. (eds.) *SUM 2015. LNCS (LNAI)*, vol. 9310, pp. 249–263. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-23540-0_17

6. Dubois, D., Prade, H.: Possibility Theory: An Approach to Computerized Processing of Uncertainty. Plenum Press, New York (1988)
7. Dubois, D., Foulloy, L., Mauris, G., Prade, H.: Probability-possibility transformations, triangular fuzzy sets, and probabilistic inequalities. *Reliable Comput.* **10**, 273–297 (2004)
8. Lee, L.: On the effectiveness of the skew divergence for statistical language analysis. In: *Proceedings of Artificial Intelligence and Statistics (AISTATS)*, pp. 65–72 (2001)
9. Heckerman, D., Breese, J.: A new look at causal independence. In: *Proceedings of the Tenth Annual Conference on Uncertainty in Artificial Intelligence (UAI-94)*, pp. 286–292, Morgan Kaufmann, San Francisco (1994)
10. Glover, F.: Future paths for integer programming and links to artificial intelligence. *Comput. Oper. Res.* **13**(5), 533–549 (1986)
11. Glover, F.: Tabu Search-Part I. *ORSA J. Comput.* **1**(3), 190–206 (1989)
12. Glover, F.: Tabu Search-Part II. *ORSA J. Comput.* **2**(1), 4–32 (1990)
13. Petiot, G.: The calculation of educational indicators by uncertain gates. In: *Proceedings of the 20th International Conference on Enterprise Information Systems*, pp. 379–386, SCITEPRESS, Funchal, Madeira (2018)
14. Petiot, G.: Merging information using uncertain gates: an application to educational indicators. In: Medina, J., et al. (eds.) *IPMU 2018. CCIS*, vol. 853, pp. 183–194. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-91473-2_16
15. Korb, K.B., Nicholson, A.E.: *Bayesian Artificial Intelligence*, 2nd edn. CRC Press, Boca Raton (2010). Section 2.2.2
16. Lauritzen, S.L., Spiegelhalter, D.J.: Local computations with probabilities on graphical structures and their application to expert systems. *J. R. Stat. Soc.: Series B (Methodol.)* **50**, 157–194 (1988)
17. Neapolitan, R.E.: *Probabilistic Reasoning in Expert Systems: Theory and Algorithms*. Wiley, New-York (1990)
18. Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E.: Equation of state calculations by fast computing machines. *J. Chem. Phys.* **21**(6), 1087–1092 (1953)
19. Scutari, M., Denis, J.-B.: *Bayesian Networks: with Examples in R*. Chapman & Hall (2014)
20. Pearl, J.: A new look at causal independence. In: *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, San Mateo (1988)
21. Zadeh, L.A.: Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets Syst.* **1**, 3–28 (1978)
22. Yager, R.R.: On ordered weighted averaging aggregation operators in multicriteria decisionmaking. *IEEE Trans. Syst. Man Cybern.* **18**(1), 183–190 (1993)
23. Sachs, K.: Causal protein-signaling networks derived from multiparameter single-cell data. *Science* **309**, 1187–1187 (2005)
24. Zagorecki, A., Druzdzal, M.: Knowledge engineering for Bayesian networks: how common are noisy-MAX distributions in practice? *IEEE Trans. Syst. Man. Cybern.: Syst.* **43**, 186–195 (2013)
25. Zagorecki, A., Druzdzal, M.: Probabilistic independence of causal influences. In: *Proceedings of the Third European Workshop on Probabilistic Graphical Models (PGM-2006)*, pp. 325–332 (2006)