Compiling Possibilistic Networks to Compute Learning Indicators

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

Keywords: Uncertain Gates, Compiling Possibilistic Networks, Possibility Theory, Education, Decision Making.

Abstract: University teachers, who generally focus their interest on pedagogy and students, may find it difficult to manage e-learning platforms which provide learning analytics and data. But learning indicators might help teachers when the amount of information to process grows exponentially. The indicators can be computed by the aggregation of data and by using teachers' knowledge which is often imprecise and uncertain. Possibility theory provides a solution to handle these drawbacks. Possibilistic networks allow us to represent the causal link between the data but they require the definition of all the parameters of Conditional Possibility Tables. Uncertain gates allow the automatic calculation of these Conditional Possibility Tables by using for example the logical combination of information. The calculation time to propagate new evidence in possibilistic networks can be improved by compiling possibilistic networks. In this paper, we will present an experimentation of compiling possibilistic networks to compute course indicators. Indeed, the LMS Moodle provides a large scale of data about learners that can be merged to provide indicators to teachers in a decision making system. Thus, teachers can propose differentiated instruction which, better corresponds to their student's expectations and their learning style.

1 INTRODUCTION

Modeling indicators based on expert knowledge are hard to perform because human description is often vague. Possibility theory, introduced by L. A. Zadeh (Zadeh, 1978), is a solution to this problem of uncertainty which appears during knowledge modeling. Moreover, the causal link between the data can be modeled by using the possibilistic network (Benferhat et al., 1999). The latter is an adaptation of the Bayesian Network (Pearl, 1988; Neapolitan, 1990) to possibility theory. In the possibilisic network each variable is attached to a Conditional Possibility Table. The number of parameters to elicit in a CPT grows exponentially depending proportionally on the number of parents. So a solution can be to use uncertain logical gates between the variables in order to compute automatically the CPT instead of eliciting all parameters. This time-saving solution allows us to fix the problem of unknown variables which are too difficult to extract from complex systems. The addition of a variable called leakage variable leads to a new model. There is a large set of available connectors from behavior severe to indulgent. The variables are often qualitative as in (Dubois et al., 2015) but to use uncertain gates we have to encode the modalities into numerical values.

Another problem is the computation time of the propagation of evidence in possibilistic networks. There are several existing solutions with exact inference or approximative inference. For example forward-chaining, message passing in junction tree, etc. But in our study we propose to experiment a new approach which is more efficient. Indeed, it is possible to perform the compiling of the possibilistic network as for Bayesian networks (Park and Darwiche, 2002) to improve the computation time.

In this paper, we would like to perform an experimentation of indicator calculation by using uncertain gates and compiling possibilistic networks. Several studies were performed in order to improve pedagogy and understand students and their learning style (Huebner, 2013; Baker and Yacef, 2009; Bousbia et al., 2010). These researchers made use of Bayesian networks, neural networks, support vector machines. They often tried to detect a student at the risk of dropping out or failing at the examination.

Our approach study is based on an existing dataset built from Moodle logs for a course of spreadsheet and some external information such as attendance and results at the examination. This dataset is anonymized. We can extract from Moodle the results of the quiz, the sources consulted, etc. The knowledge about the indicators is provided by teachers and

Petiot, G.
Compiling Possibilistic Networks to Compute Learning Indicators.
DOI: 10.5220/0010238001690176
In *Proceedings of the 13th International Conference on Agents and Artificial Intelligence (ICAART 2021) - Volume 2*, pages 169-176
ISBN: 978-989-758-484-8

extracted from the data by data mining as in (Petiot, 2018).

The goal of this paper, is to compute course indicators by using teachers' knowledge. To do this, we will first present possibility theory and uncertain gates. Then we will focus on the compiling of possibilistic networks and finally we will present our results.

2 UNCERTAIN GATES

Uncertain gates are an analogy of noisy gates in possibility theory, developed in 1978 by L.A. Zadeh (Zadeh, 1978). In this theory, imprecise and uncertain knowledge can be modeled by a possibility distribution π. We can define the possibility measure Π and the necessity measure N from $P(\Omega)$ in $[0,1]$ as the authors in (Dubois and Prade, 1988). The possibility measure is defined as follows:

$$\forall A \in P(\Omega), \Pi(A) = \sup_{x \in A} \pi(x) \quad (1)$$

The necessity measure can be defined as follows:

$$\forall A \in P(\Omega), N(A) = 1 - \Pi(\neg A) = \inf_{x \notin A} 1 - \pi(x) \quad (2)$$

Possibility theory is not additive but maxitive:

$$\forall A, B \in P(X), \Pi(A \cup B) = \max(\Pi(A), \Pi(B)). \quad (3)$$

We can compute the possibility of the variable A given the variable B by using the conditioning proposed by E. Hisdal (Hisdal, 1978) and generalized by D. Dubois and H. Prade (Dubois and Prade, 1988):

$$\Pi(A|B) = \begin{cases} \Pi(A,B) & \text{if } \Pi(A,B) < \Pi(B), \\ 1 & \text{if } \Pi(A,B) = \Pi(B). \end{cases} \quad (4)$$

Possibilistic networks (Benferhat et al., 1999; Borgelt et al., 2000) can be defined by using the factoring property. We propose the following definition:

Definition 2.1. The factoring property can be defined from the joint possibility distribution $\Pi(V)$ for a Directional Acyclic Graph (DAG) $G = (V, E)$ where V is the set of Variables and E the set of edges between the variables. $\Pi(V)$ can be factorized toward the graph G:

$$\Pi(V) = \bigotimes_{X \in V} \Pi(X/Pa(X)). \quad (5)$$

The function $Pa(X)$ returns the parents of the variable X.

There are two kinds of possibilistic networks: min-based possibilistic networks that are qualitative possibilistic networks where $\bigotimes$ is the function min, and product-based possibilistic networks that are quantitative possibilistic networks where $\bigotimes$ is the product. In this research, we will use a min-based possibilistic network because we have chosen to compare the possibilistic values instead of using an intensity scale in [0,1].

Uncertain logical gates were proposed for the first time by the authors of (Dubois et al., 2015). They are based on the Independence of Causal Influence and use a model to represent uncertainty. This model is built by introducing an intermediate variable Z_i between a set of causal variables $X_1, ..., X_n$ and an effect variable Y. This allows us to represent two behaviors: inhibitors and substitute. The inhibitors can be defined if a cause is met and the effect variable Y is not produced. The substitute can be defined if a cause is not met and the variable Y is produced. In this model, there is a deterministic function f which combines the influences of the variables $Z_i s$: $Y = f(Z_1, ..., Z_n)$. The leaky model is derived from the previous model by adding a leakage variable Z_l which represents the unknown knowledge. The possibilistic model with the ICI is the following:

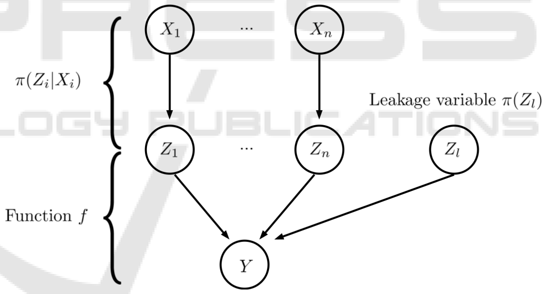


Figure 1: Possibilistic model with ICI.

This model, presented by the authors (Dubois et al., 2015), leads to the following equation:

$$\pi(y|x_1, ..., x_n) = \max_{z_1, ..., z_n, z_l : y = f(z_1, ..., z_n, z_l)} \bigotimes_{i=1}^{n} \pi(z_i|x_i) \otimes \pi(z_l) \quad (6)$$

The $\otimes$ is the minimum and $\oplus$ is the maximum. There are several possible functions f, for example AND, OR, NOT, INV, XOR, MAX, MIN, MEAN, linear combination, etc.

In order to generate the CPT we have to compute the above equation. We have to define $\pi(Z_i|X_i)$, $\pi(Z_l)$, and the function f. In our experimentation we have three ordered levels of intensity: low, medium and high. We propose to encode the modality by the following intensity levels as in (Dubois et al., 2015): 0 for low, 1 for medium and 2 for high. The following table illustrates an example:

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 above 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 $\pi(Z_l)$. Another constraint is that $\kappa_i^{1,2} \geq \kappa_i^{0,2}$.

In our study we will use for the function f the function MIN and MAX leading to the connectors uncertain MIN ($\perp$) and uncertain MAX ($\top$) as proposed by Dubois et al. (Dubois et al., 2015). We will also use a weighted average function (WAVG) and a MYCIN Like connector ($\hbar$) as described in (Petiot, 2018). 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. Nevertheless, the weighted average function can return a value outside the domain of Y. We propose to combine the result of the weighted average function $g(z_1,...,z_n) = \omega_1 z_1 + ... + \omega_n z_n$ with a scaling function f_s which returns a value in the domain of Y. The parameters ω_i are the weights of the weighted average. Finally, we have $f = f_s \circ g$. If $(\varepsilon_0, \varepsilon_1, ..., \varepsilon_{m-1})$ are the m ordered states of Y then the function f_s can be for example:

$$f_s(x) = \begin{cases} \varepsilon_0 & \text{if } x \leq \theta_0 \\ \varepsilon_1 & \text{if } \theta_0 < x \leq \theta_1 \\ \vdots & \vdots \\ \varepsilon_{m-1} & \text{if } \theta_{m-2} < x \end{cases} \quad (7)$$

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 (connector $\sum$).

3 COMPILING THE JUNCTION TREE OF A POSSIBILISTIC NETWORK

The knowledge of the indicators is represented by a possibilistic network. The propagation of evidence in this possibilistic network will lead to a possibility and a certainty measure.

To perform the propagation of evidence in the possibilistic network we propose at first to compile the junction tree of a possibilistic network. The junction tree is composed of cliques and separators. The cliques are extracted by using the Kruskal algorithm (Kruskal, 1956) after the generation of the moral graph and the triangulated graph (Kjaerulff, 1994). The same reasoning as in compiling Bayesian networks (Darwiche, 2003) is used. We have adapted A. Darwiche's algorithm for the junction tree (Park and Darwiche, 2002) of a Bayesian network to possibilistic networks.

Indeed, possibilistic networks can be transformed into a function with two kind of variables: evidence indicators and network parameters. For all instantiations of a variable $X = x$ we define an evidence indicator λ_x. Similarly, for all network CPT parameters of $\pi(X|U)$, we define a parameter $\theta_{x|u}$ where u is an instantiation of U, the parents of the variable X and x an instantiation of the variable X.

The function f can be computed by combining at first all evidence indicators and network parameters consistent with the instantiation by using the operator $\otimes$. Then, we perform a combination of all the previous results by using the operator $\oplus$.

Definition 3.1. If P is a possibilistic network, $V = v$ the instantiations of the variables of the possibilistic networks and $U = u$ the consistent instantiations of the parents of a variable X with the instantiation $X = x$, then the function f of P is:

$$f = \bigoplus_{v} \bigotimes_{xu \sim v} \lambda_x \otimes \theta_{x|u} \quad (8)$$

In the above formula xu denotes the instantiation of the family of X and its parents U compatible with the instantiation v. The operator $\bigoplus$ can be the function maximum and $\bigotimes$ the function minimum.

We can study, as an example, the following possibilistic network:

Table 2: Example of the possibilistic network $A \rightarrow B \rightarrow C$.

A	B	
true	true	$\theta_{b\|a}$
true	false	$\theta_{\bar{b}\|a}$
false	true	$\theta_{b\|\bar{a}}$
false	false	$\theta_{\bar{b}\|\bar{a}}$

A	
true	θ_a
false	$\theta_{\bar{a}}$

B	C	
true	true	$\theta_{c\|b}$
true	false	$\theta_{\bar{c}\|b}$
false	true	$\theta_{c\|\bar{b}}$
false	false	$\theta_{\bar{c}\|\bar{b}}$

In this case the function f is:

$$f = \lambda_a \otimes \lambda_b \otimes \lambda_c \otimes \theta_a \otimes \theta_{b|a} \otimes \theta_{c|b} \\ \oplus \lambda_a \otimes \lambda_b \otimes \lambda_{\bar{c}} \otimes \theta_a \otimes \theta_{b|a} \otimes \theta_{\bar{c}|b} \\ \vdots \\ \oplus \lambda_{\bar{a}} \otimes \lambda_{\bar{b}} \otimes \lambda_{\bar{c}} \otimes \theta_{\bar{a}} \otimes \theta_{\bar{b}|\bar{a}} \otimes \theta_{\bar{c}|\bar{b}} \tag{9}$$

Definition 3.2. If the evidence e is an instantiation of variables then we have the property $f(e) = \pi(e)$.

Let us consider the following example:

Table 3: Example of the possibilistic network $A \to B$.

A	B	
true	true	1
true	false	0.2
false	true	0.1
false	false	1

A	
true	1
false	0.1

If the evidence is $\bar{a}$, then we obtain $\lambda_a = 0$, $\lambda_{\bar{a}} = 1$, $\lambda_b = 1$, $\lambda_{\bar{b}} = 1$ and the computation of $f(e)$ is: $f(\bar{a}) = f(\lambda_a = 0, \lambda_{\bar{a}} = 1, \lambda_b = 1, \lambda_{\bar{b}} = 1) = \theta_{\bar{a}} \otimes \theta_{b|\bar{a}} \oplus \theta_{\bar{a}} \otimes \theta_{\bar{b}|\bar{a}} = 0.1 \otimes 0.1 \oplus 0.1 \otimes 1.0 = 0.1$. The evaluation of f leads us to compute $\pi(e)$.

We can compute the possibility of a variable X given the evidence e:

$$\pi(x|e) = \begin{cases} \pi(x,e) & \text{if } \pi(x,e) < \pi(e), \\ 1 & \text{if } \pi(x,e) = \pi(e). \end{cases} \tag{10}$$

If the variable X has n states and x is one of its states, and if X is not in the evidence e then $\pi(x,e) = f(e, 1_{\lambda_x})$ with $1_{\lambda_x} = (\lambda_{x_1} = 0, ..., \lambda_x = 1, ..., \lambda_{x_n} = 0)$. For example, $1_{\lambda_{\bar{a}}} = (\lambda_a = 0, \lambda_{\bar{a}} = 1)$. We have another case if X is in e called Evidence Retraction in probability theory (Darwiche, 2003). In possibility theory the Evidence Retraction leads us to compute $\pi(x|e - X)$.

If the number of variables is too high, the computing of possibilities becomes too complex. So it is interesting to compile the possibilistic network by using a MIN-MAX circuit as in (Raouia et al., 2010). This optimization allows us to reduce memory used and computation time. There are several approaches for compiling the function f into a MIN-MAX circuit. The leaves of the MIN-MAX circuit are the parameters λ and θ and the nodes are the operators $\otimes$ and $\oplus$. We present an example of a MIN-MAX circuit in the following figure:

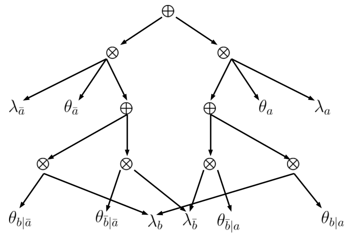
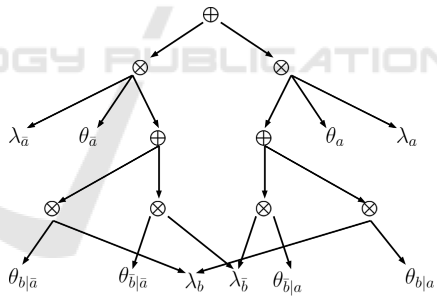


Figure 2: MIN-MAX circuit of the example.

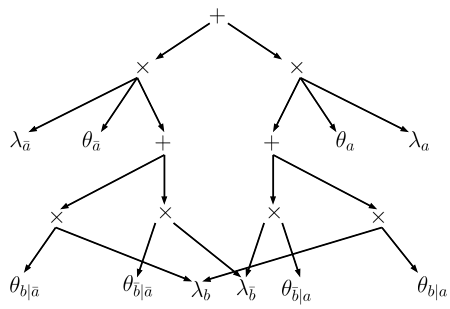
We have chosen to perform the factorisation of the function f and then to use the junction tree method. To compile a Bayesian network under evidence, we generate an arithmetic circuit and we differentiate the circuit in order to obtain all posterior probabilities $p(x|e)$. The differentiation is very easy with the arithmetic circuit in probability theory.

Definition 3.3. We obtain the arithmetic circuit f' of a MIN-MAX circuit f by replacing the $\otimes$ by the multiplications and the $\oplus$ by additions.

We propose to encode the MIN-MAX circuit into an arithmetic circuit. Then we will deduce the propagation algorithm for the arithmetic circuit. Finally, we will replace in the algorithm the additions by $\oplus$ and the multiplications by $\otimes$ in order to apply the algorithm to a MIN-MAX circuit. In the following figure we present an example of our encoding:



(a) MIN-MAX circuit.



(b) Arithmetic circuit.

Figure 3: Arithmetic circuit of a MIN-MAX circuit.

We can differentiate the polynomial $\frac{\partial f'}{\partial \lambda_x}$. In the example of Table 3, we obtain the following function:

$$f = \lambda_a \otimes \theta_a \otimes (\lambda_b \otimes \theta_{b|a} \oplus \lambda_{\bar{b}} \otimes \theta_{\bar{b}|a}) \\ \oplus \lambda_{\bar{a}} \otimes \theta_{\bar{a}} \otimes (\lambda_b \otimes \theta_{b|\bar{a}} \oplus \lambda_{\bar{b}} \otimes \theta_{\bar{b}|\bar{a}}) \tag{11}$$

After the transformation we obtain the following polynomial:

$$f' = \lambda_a \theta_a (\lambda_b \theta_{b|a} + \lambda_{\bar{b}} \theta_{\bar{b}|a}) + \lambda_{\bar{a}} \theta_{\bar{a}} (\lambda_b \theta_{b|\bar{a}} + \lambda_{\bar{b}} \theta_{\bar{b}|\bar{a}}) \tag{12}$$

For example, if we suppose that $e = b$ then $f'(e) = f'(\lambda_b = 1; \lambda_{\bar{b}} = 0; \lambda_a = 1; \lambda_{\bar{a}} = 1)$. To compute $\pi(a, e)$ we must at first compute $\frac{\partial f'(e)}{\partial \lambda_a}$ because a is not in e. We obtain the following result:

$$\frac{\partial f'(e)}{\partial \lambda_a} = \theta_a \theta_{b|a} \tag{13}$$

To obtain $\pi(a, e)$ we must replace the additions by $\oplus$ and the multiplications by $\otimes$ in the above equation, which gives the following result:

$$\pi(a, e) = \theta_a \otimes \theta_{b|a} \tag{14}$$

We propose to build the MIN-MAX circuit of a junction tree obtained from a possibilistic network. We must first select a root node which is the result of f, then we add a $\oplus$ node for each instantiation of a separator and a $\otimes$ node for each instantiation of a cluster. We have only one node λ_x for each instantiation of a variable X and one node $\theta_{x|v}$ for each instantiation of the nodes X and its parents V. The children of the output node f are the $\otimes$ nodes of the root cluster. The children of $\oplus$ nodes are compatible nodes generated by the child clusters and the children of a $\otimes$ node are compatible nodes generated by the child separators. Each variable and λ are affected to only one cluster.

If we consider the example $B \longleftarrow A \longrightarrow C$ we can compute the MIN-MAX circuit of the junction tree as follows:

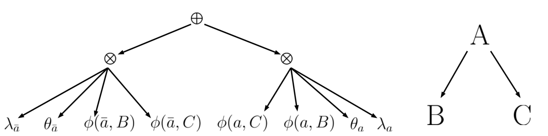


Figure 4: MIN-MAX circuit of a junction tree.

In this figure, the function ϕ performs the evaluation of the cluster compatible with the instantiation of the separator.

We propose now to differentiate the arithmetic circuit f' of a MIN-MAX circuit f. If v is the current node and p is the parents of v, then we can compute $\frac{\partial f'}{\partial v}$ by using the chain rule:

$$\frac{\partial f'}{\partial v} = \sum_p \frac{\partial f'}{\partial p} \frac{\partial p}{\partial v} \tag{15}$$

If the parent p has n other children v_i different from the node v, there are several cases to discuss:

- If v is the first node then $\frac{\partial f'}{\partial v} = 1$.
- If p is an addition node then $\frac{\partial p}{\partial v} = \frac{\partial (v + \sum_{i=1}^{n} v_i)}{\partial v} = 1$
- If p is a multiplication node then $\frac{\partial p}{\partial v} = \frac{\partial (v \prod_{i=1}^{n} v_i)}{\partial v} = \prod_{i=1}^{n} v_i$

As a result, we obtain the following recursive algorithm to evaluate the MIN-MAX circuit of a junction tree by changing the multiplication by $\otimes$ and additions by $\oplus$:

1. Upward-pass: compute the value of the node v and store it in $u(v)$;
2. If v is the root then set $d(v) = 1$ else set $d(v) = 0$;
3. Downward-pass: for each parent p of the node v compute $d(v)$ as follows:

(a) if p is a node $\oplus$:

$$d(v) = d(v) \bigoplus d(p) \tag{16}$$

(b) if p is a node $\otimes$:

$$d(v) = d(v) \bigoplus \left[d(p) \bigotimes \left[\bigotimes_{i=1}^{n} u(v_i^p) \right] \right] \tag{17}$$

The nodes v_i^p are the other children of p;

To evaluate the indicators, we must perform several processing operations. The first one is to compile the junction tree of the possibilistic network. Then we perform the initialization of evidence before applying the recursive algorithm. We can compute for each state of an indicator a possibility measure and a necessity measure.

4 EXPERIMENTATION

4.1 Presentation

In our experimentation, we used an existing anonymized dataset for a Spreadsheet course at bachelor level proposed in face-to-face learning enriched by an online supplement on Moodle. This dataset was built by gathering all data of logs in a table. Then a process of anonymization was performed. For example, we use the data of Moodle, such as quiz results, sources consulted, ... and external data such as attendance, groups,... The quiz questions were categorized by skills. When the data were missing, we performed an imputation of these data by an iterative PCA (Audigier et al., 2015). The knowledge about the indicators was provided by teachers and extracted from the data by data mining. To represent the knowledge we have chosen to use a DAG:

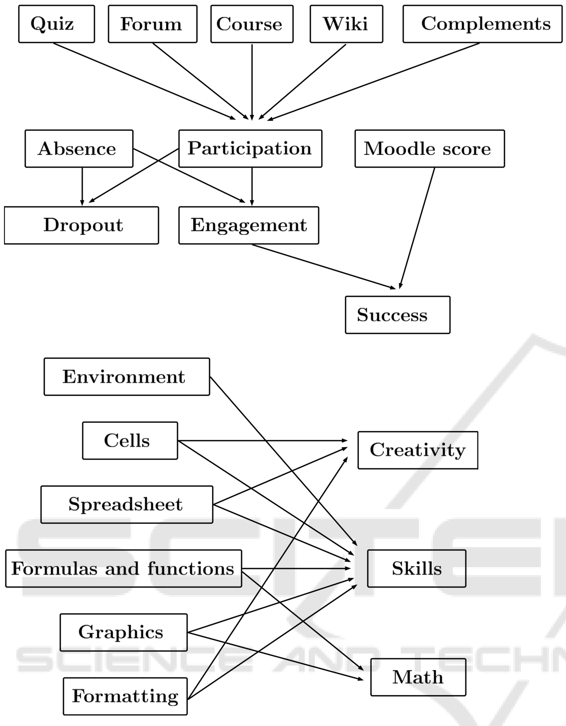
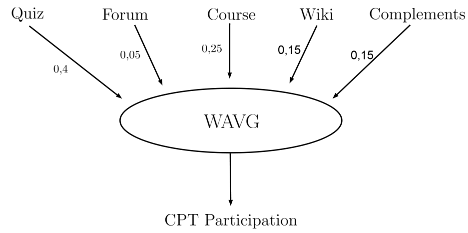


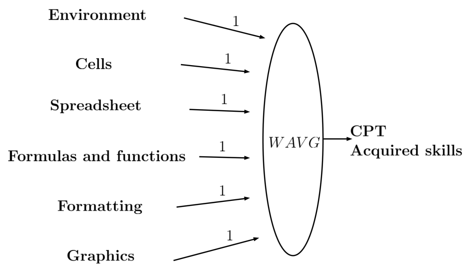
Figure 5: Modeling of knowledge by a DAG.

The qualitative variables have 3 ordered modalities (low, medium, high) encoded with the numerical values (0,1,2). The description of the indicators by teachers is often imprecise so we used a possibility distribution to represent each state of a variable. Then, possibilistic networks can be used to compute the indicators. To do this we need to define all CPTs. To avoid the eliciting of all the parameters, we used uncertain gates leading to the computation of all the CPTs.

We merged information about the sources consulted in Moodle to build an indicator of participation which takes into account their importance. We used the WAVG connector. The weights were provided by teachers. We also computed an indicator of acquired skills by using the WAVG connector. The name of this connector is connector $\sum$. We present in the following figure the weights of the WAVG connectors:



(a) Indicator of participation.



(b) Indicator of acquired skills.

Figure 6: Weights of the WAVG connectors.

We used the uncertain MIN connector ($\perp$) for conjunctive behavior and the uncertain hybrid connector ($\hbar$) for indicators which need a compromise in case of conflict and a reinforcement if the values are concordant. As a result we obtain the following model:

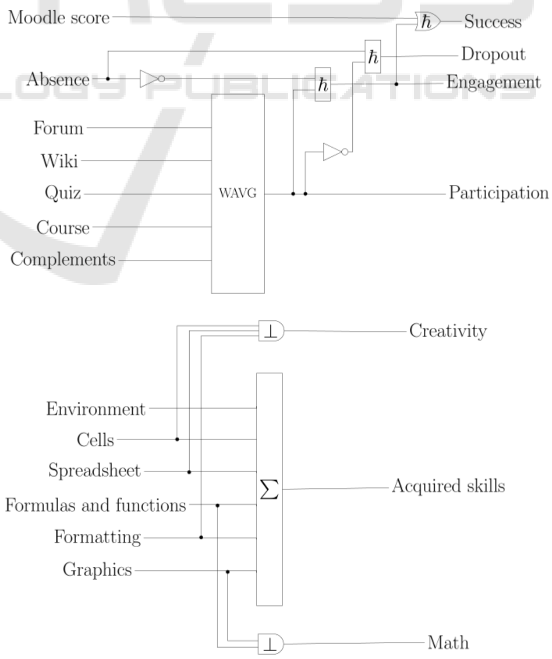
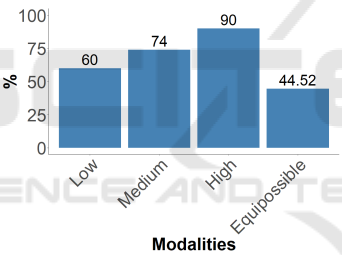


Figure 7: Knowledge modeling with uncertain connectors.

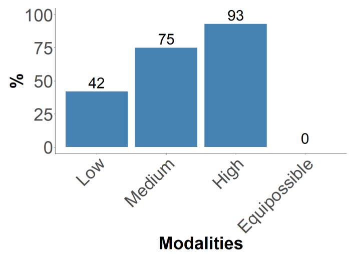
Before the propagation of the new information, we have to build the CPTs of all the uncertain gates. Then, we apply the algorithm for compiling the junction tree of a possibilistic network. We have compared this approach to a previous study that used the message passing algorithm (Petiot, 2018). Indeed, we can adapt the junction tree message passing algorithm (Lauritzen and Spiegelhalter, 1988) of Bayesian Networks to Possibilistic Networks. The propagation algorithm can be resumed in three steps. The initialization with the injection of evidence, then, the collect with the propagation of evidence from leaf to root and the distribution with the propagation of evidence from root to leaf.

4.2 Results

We have compared the compiling of possibilistic network and the message passing algorithm. As expected, the results of the indicators in both approaches are identical. For example, the indicator of success deals with the prediction of student success at the exam. We have computed the percentage of success for each state of the indicator of success. We obtain the following results by using the compiling of possibilistic networks:



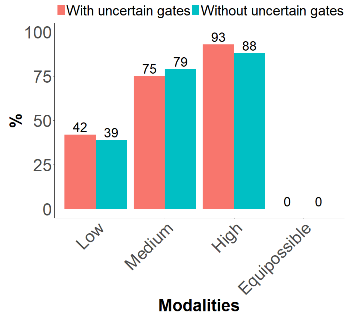
(a) Without the estimation of missing data.



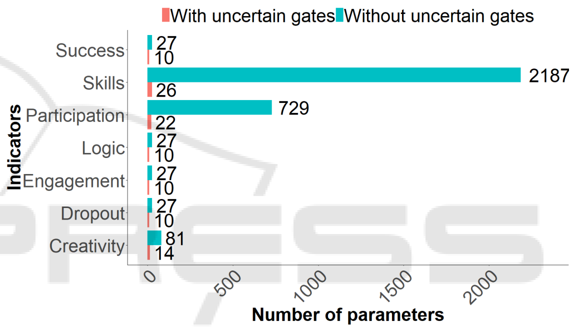
(b) With the estimation of missing data.

Figure 8: Indicator of success with and without the estimation of missing data.

We can see in figure a) a lot of equipossible results (with all possibilities equal to 1) due to missing data. To reduce the equipossible variables, we have performed an imputation of missing data using an iterative PCA algorithm (Audigier et al., 2015). We present the results in figure b). Another advantage of our approach is the use of uncertain gates in order to avoid the eliciting of all parameters of the CPTs. We have compared the result of the indicator of success with and without uncertain gates by using the compiling of the possibilistic network. The results are the following:



(a) Indicator of success.



(b) The number of parameters.

Figure 9: Comparison of the results with and without uncertain gates by using the compiling of the possibilistic network.

In figure a) the results are very close but uncertain gates require fewer parameters than CPTs elicited by a human expert. Figure b) shows that the number of parameters is highly decreased by using uncertain gates. We have also compared the performance of the computation of the indicators by using the compiling of possibilistic networks and the message passing algorithm. The results are the following:

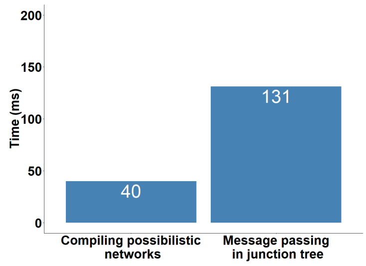


Figure 10: The mean computation time.

We can see that the computation time is improved by compiling the junction tree of the possibilistic network. The compiling approach is three times faster than the message passing algorithm.

5 CONCLUSION

In this paper, we have presented a new approach of exact inference based on the compiling of the junction tree of a possibilistic network. We applied this approach to computing learning indicators for a course of spreadsheet that can be presented in a decision making system for teachers. To do this we have represented teachers' knowledge by using a possibilistic network. As the number of parameters of the CPT grows exponentially when the number of parents grows, we have proposed to use uncertain gates because they allows us to avoid eliciting all CPT parameters. The CPTs are computed automatically. Then, we have computed the junction tree and generated the MIN-MAX circuit. To compute the possibilities of the indicators we have applied our algorithm which begins by an upward pass followed by a downward pass. We have shown that the computation time is improved compared to our previous inference approach based on the message passing algorithm. The results of our approach and message passing algorithm were the same as expected. In future, we would like to perform further experimentations in order to better evaluate our junction tree compiling approach for possibilistic networks. We would like to perform further experimentation concerning the computation of learning indicators.

REFERENCES

Audigier, V., Husson, F., and Josse, J. (2015). A principal components method to impute missing values for mixed data. *Advances in Data Analysis and Classification*, 10(1):5–26.

Baker, R. S. J. d. and Yacef, K. (2009). The state of educational data mining in 2009: A review and future visions. *Journal of Educational Data Mining*, 1(1):3–17.

Benferhat, S., Dubois, D., Garcia, L., and Prade, H. (1999). Possibilistic logic bases and possibilistic graphs. *In Proc. of the Conference on Uncertainty in Artificial Intelligence*, pages 57–64.

Borgelt, C., Gebhardt, J., and Kruse, R. (2000). Possibilistic graphical models. *Computational Intelligence in Data Mining*, 26:51–68.

Bousbia, N., Labat, J. M., Balla, A., and Rebaï, I. (2010). Analyzing learning styles using behavioral indicators in web based learning environments. *EDM 2010 International Conference on Educational Data Mining*, pages 279–280.

Darwiche, A. (2003). A differential approach to inference in bayesian networks. *J. ACM*, 50(3):280–305.

Dubois, D., Fusco, G., Prade, H., and Tettamanzi, A. G. B. (2015). Uncertain logical gates in possibilistic networks. an application to human geography. *Scalable Uncertainty Management 2015*, pages 249–263.

Dubois, D. and Prade, H. (1988). *Possibility theory: An Approach to Computerized Processing of Uncertainty*. Plenum Press, New York.

Hisdal, E. (1978). Conditional possibilities independence and noninteraction. *Fuzzy Sets and Systems*, 1(4):283 – 297.

Huebner, R. A. (2013). A survey of educational data mining research. *Research in Higher Education Journal*, 19.

Kjaerulff, U. (1994). Reduction of computational complexity in bayesian networks through removal of week dependences. *Proceeding of the 10th Conference on Uncertainty in Artificial Intelligence*, pages 374–382.

Kruskal, J. B. (1956). On the shortest spanning subtree of a graph and the travelling salesman problem. *Proceedings of the American Mathematical Society*, pages 48–50.

Lauritzen, S. and Spiegelhalter, D. (1988). Local computation with probabilities on graphical structures and their application to expert systems. *Journal of the Royal Statistical Society*, 50(2):157–224.

Neapolitan, R. E. (1990). *Probabilistic reasoning in expert systems: theory and algorithms*. John Wiley & Sons.

Park, J. D. and Darwiche, A. (2002). A differential semantics for jointree algorithms. In *Advances in Neural Information Processing Systems 15 [Neural Information Processing Systems, NIPS 2002, December 9-14, 2002, Vancouver, British Columbia, Canada]*, pages 785–784.

Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.

Petiot, G. (2018). Merging information using uncertain gates: An application to educational indicators. 853:183–194.

Raouia, A., Amor, N. B., Benferhat, S., and Rolf, H. (2010). Compiling possibilistic networks: Alternative approaches to possibilistic inference. In *Proceedings of the Twenty-Sixth Conference on Uncertainty in Artificial Intelligence*, UAI'10, pages 40–47, Arlington, Virginia, United States. AUAI Press.

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