

Intégration de l'Evidence Floue dans les Réseaux Bayésiens

Ali Ben Mrad

Université de Sfax, ENIS
ENIS, CES
Sfax, Tunisia
benmradali2@gmail.com

Mohamed Amine Maalej

Université de Sfax, FSS
ENIS, CES
Sfax, Tunisia
m.maalej@gmail.com

Veronique Delcroix

Université Lille Nord de France, F-59000, Lille, France
UVHC, LAMIH FRE CNRS 3304, F-59313,
Valenciennes, France
Veronique.Delcroix@univ-valenciennes.fr

Sylvain Piechowiak

Université Lille Nord de France, F-59000, Lille, France
UVHC, LAMIH FRE CNRS 3304, F-59313,
Valenciennes, France
Sylvain.Piechowiak@univ-valenciennes.fr

Mohamed Abid

Université de Sfax, ENIS
ENIS, CES
Sfax, Tunisia
mohamed.abid_ces@yahoo.fr

Abstract— Bayesian Networks are graphical models that are easy to interpret and update. These models are useful if the knowledge is uncertain, but they lack some means to express ambiguity. To face this problem, we propose Fuzzy Evidence in Bayesian Networks and combine the Fuzzy Logic and Bayesian Network. This has allowed to benefit from mutual advantages of these two approaches, and to overcome the problem of data and observation ambiguity. This paper proposes an inference algorithm which uses the Bayesian Network and Fuzzy Logic reliability. This solution has been implemented, tested and evaluated in comparison with the existing methods.

Keywords- Bayesian Networks, Fuzzy Logic, Uncertain Evidence, Artificial Intelligence.

Résumé— Les Réseaux Bayésiens sont des modèles graphiques performants pour la manipulation de connaissance incertaine. Cependant, l'une des limites majeures que rencontrent les utilisateurs des Réseaux Bayésiens est d'exprimer les connaissances floues.

Face à ce problème, nous proposons l'intégration de l'évidence floue dans les Réseaux Bayésiens en combinant la logique floue avec les Réseaux Bayésiens. Ceci permet de profiter des avantages mutuels de ces deux approches et de surmonter le problème rencontré à savoir l'ambiguïté des données.

Cet article propose un algorithme d'inférence pour l'évidence floue dans les Réseaux Bayésiens. Cet algorithme a été implémenté puis testé et évalué par rapport aux méthodes existantes.

Mots clés- Réseaux Bayésiens, Logique Floue, Evidence Incertaine, Intelligence Artificielle.

I. INTRODUCTION

Le domaine de la gestion des connaissances, qui connaît un intérêt croissant, et donc également un champ d'application potentiel pour les Réseaux Bayésiens, qui sont actuellement l'une des techniques les plus intéressantes de l'Intelligence Artificielle, dans la mesure où ceux-ci offrent un formalisme riche et intuitif de représentation de la connaissance.

En effet, les Réseaux Bayésiens (RBs) [1, 2] constituent aujourd'hui l'un des formalismes les plus complets et les plus cohérents pour l'acquisition, la représentation et l'utilisation de connaissances par ordinateurs.

Dans certains cas, l'utilisation des Réseaux Bayésiens [3] se trouve limitée par l'impossibilité d'utiliser des variables de natures flous.

La logique floue est une approche développée par Zadeh [4], basée sur sa théorie des sous-ensembles flous (fuzzy sets en anglais), généralisant la théorie des ensembles classiques.

Dans la nouvelle théorie de Zadeh, un élément peut plus ou moins appartenir à un certain ensemble. Les imprécisions et les incertitudes peuvent ainsi être modélisées, et les raisonnements acquièrent une flexibilité que ne permet pas la logique classique.

Cet article décrit la possibilité d'intégrer l'évidence floue dans les Réseaux Bayésiens.

L'objectif de ce travail est d'appréhender la problématique des Réseaux Bayésiens avec évidence incertaine, ensuite de modéliser quelques systèmes par des

Réseaux Bayésiens avec évidence floue et de développer des algorithmes d'inférence à des fins de diagnostic exploitant la richesse de ces modèles.

II. RÉSEAU BAYÉSIEN

A. Définition

Classiquement, un Réseau Bayésien est un outil de représentation des connaissances, qui permet de calculer des probabilités conditionnelles, apportant ainsi des solutions à différentes sortes de problématiques. La structure de ce type de réseau est simple : En effet, selon Judea Pearl « *les réseaux bayésiens sont des graphes acycliques orientés pour lesquels les nœuds représentent des variables aléatoires et les arcs représentent l'indépendance conditionnelle entre les différents nœuds* » [5].

Formellement, un réseau bayésien est défini par Naïm et al dans [1] par :

- un graphe acyclique orienté $G, G = (V, E)$, où V est l'ensemble des nœuds de G , et E l'ensemble des arcs de G ,
- un espace probabiliste fini (Ω, Z, p) ,
- un ensemble de variables aléatoires associées aux nœuds du graphe et définies sur (Ω, Z, p) tel que :

$$p(V_1, V_2, \dots, V_n) = \prod_{i=1}^n p(V_i | C(V_i)), \text{ où } C(V_i)$$

est l'ensemble des causes (parents) de V_i dans le graphe G .

Cependant, les réseaux bayésiens doivent leur nom aux travaux de Thomas Bayes au XVIII^e siècle sur la théorie de probabilités, par son célèbre théorème.

$$P(M_i | A) = \frac{P(A | M_i)P(M_i)}{\sum_k P(A | M_k)P(M_k)}$$

$P(M_i)$: probabilité *a priori* de M_i .

$P(A | M_i)$: probabilité de A conditionnellement à M_i .

$P(M_i | A)$: probabilité *a posteriori* de M_i conditionnellement à A .

L'inférence, ou *la mise à jour des probabilités*, dans un réseau bayésien se résume à un calcul de probabilités *a posteriori*. Connaissant les états de certaines variables (appelées variables d'observation), on détermine les probabilités des états de certaines autres variables (appelées variables cibles) conditionnellement aux observations. Parmi les algorithmes d'inférence les plus utilisés *l'arbre de jonction*.

B. Algorithme arbre de jonction

Cet algorithme est applicable dans tous types de réseaux, arbre ou non arbre. En premier lieu il transforme le graphe en un arbre de jonction, puis il initialise ses potentiels, ensuite il utilise la méthode message passing pour la propagation des messages et le calcul des probabilités *a posteriori*.

En effet, l'algorithme se comporte de la façon suivante [6] :

- La phase de construction (ou transformation du graphe) : elle nécessite un ensemble de sous-étapes permettant de transformer le graphe initial en un arbre de jonction, dont les nœuds sont des cliques (regroupement) de nœuds du graphe initial. Cette transformation est nécessaire, d'une part pour éliminer les boucles du graphe, et d'autre part, pour obtenir un graphe plus efficace quant au temps de calcul nécessaire à l'inférence, mais qui reste équivalent au niveau de la distribution de probabilité représentée. Cette transformation se fait en trois étapes : moralisation, triangulation et création de l'arbre de jonction.
- La phase d'initialisation : il s'agit d'initialiser les potentiels des cliques et séparateurs.
- La phase de propagation : il s'agit de la phase de calcul probabiliste à proprement parler où les nouvelles informations concernant une ou plusieurs variables sont propagées dans l'ensemble du réseau, de manière à mettre à jour l'ensemble des distributions de probabilités du réseau. Ceci se fait en passant des messages contenant une information de mise à jour entre les cliques de l'arbre de jonction précédemment construit.
- La phase de marginalisation : l'arbre de jonction contiendra la distribution de probabilité sachant les nouvelles informations, c'est-à-dire $p(U | e)$ où U représente l'ensemble des variables du réseau bayésien et e l'ensemble des nouvelles informations sur les variables.

L'évidence classique est une observation d'une variable aléatoire ayant une valeur particulière. Cependant, il n'est pas toujours possible d'observer la valeur exacte d'une variable ou d'avoir une confiance complète sur une observation, d'où l'apparition de l'évidence incertaine.

III. EVIDENCE INCERTAINE

L'observation classique dans les RBs touche seulement l'un des états du nœud observé, autrement dit, si nous avons une observation dans un nœud X , seulement l'un de ses états sera observé. Ainsi, nous associons la valeur 1 à cet état et 0 pour les autres.

Dans ce cas, nous ne pouvons pas représenter le cas où notre observation est incertaine et concerne en même temps plusieurs états selon des pourcentages bien définis.

Généralement, il y a deux types d'évidences incertaines, l'évidence virtuelle qui peut être interprétée comme évidence avec incertitude et représentée par des ratios de probabilité et soft évidence interprétée comme évidence de l'incertitude et représentée par des distributions de probabilité.

Chacun de ces deux types d'évidence possède ses propres caractéristiques et obéit à une mise à jour de croyance qui diffère l'une de l'autre et diffère aussi de celle de l'évidence classique.

A. Evidence Virtuelle

L'Evidence Virtuelle (EV) a été introduite par Pearl [5] et définie comme une généralisation de l'évidence standard dans les réseaux bayésiens. Ainsi, elle propose l'intégration des connaissances externes du réseau bayésien comme indiqué dans [6, 7].

Bilmes et Reynolds ont démontré dans [8, 9], que l'EV peut augmenter la puissance de modélisation de réseaux bayésiens de manière significative sans compliquer la méthodologie d'inférence fondamentale.

L'EV est utile dans plusieurs cas par exemple il y a des scénarios où on peut raisonner seulement sur des rapports de vraisemblances plutôt que sur les valeurs de vraisemblances eux-mêmes.

Pearl a proposé dans [10] la méthode des évidences virtuelles pour traiter la mise à jour de croyance dans les RBs lorsqu'on a un événement incertain. Par exemple, supposons que nous considérons des informations externes influant sur l'état (a) d'une variable (X_i) du RB.

Ainsi, l'EV consiste à créer un nœud binaire virtuel (U) avec un état (u) qui influence sur la probabilité de la réalisation de l'événement ($X_i = a$).

Le nœud virtuel (U) possède un seul parent X_i et sa table de probabilité conditionnelle satisfait $L(X_i) = P(u|X_i=a) : P(u|X_i \neq a)$.

Dans cette méthode qui est généralisée dans [11], l'évidence virtuelle sur $Y \subseteq X$ est représentée par le rapport de vraisemblance suivant :

$$L(Y) = P(\text{ob}(y_{(1)})|y_{(1)}) : P(\text{ob}(y_{(2)})|y_{(2)}) : \dots : P(\text{ob}(y_{(n)})|y_{(n)}),$$

Avec $y_{(1)}, y_{(2)}, \dots, y_{(n)} \in Y$ représentent les instances de Y, $\text{ob}(y_{(i)})$ définit l'événement observé $Y = y_{(i)}$ est vrai, et $P(\text{ob}(y_{(i)})|y_{(i)})$ est la probabilité de l'observation de $Y = y_{(i)}$ sachant que Y est à l'état $y_{(i)}$.

Prenons un exemple de RB détaillé dans [5]. Ainsi, nous avons initialement trois variables C pour cambriolage, S pour alarme sonore déclenchée et W pour un témoignage de Watson comme indique la figure 1.

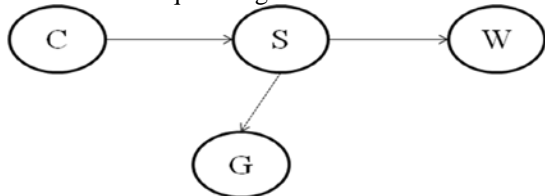


Figure 1. Exemple de RB avec Evidence Virtuelle

Dans cet exemple, notre but est de calculer la probabilité de C (un cambriolage survenu) à partir des connaissances données sur les éléments alarme déclenchée $S = 1$ et témoignage de Watson W.

Cependant, dans cet exemple, on ne sait pas si l'alarme s'est déclenchée, ainsi la variable S sera cachée. La seule chose qui est connue est peut-être quelques informations extérieures du RB influant sur la valeur de S.

Dans notre cas, M. Holmes, qui est un observateur externe, nous fournit un jugement mental, basée sur le témoignage de Mme Gibbons, ce qui influe sur l'hypothèse « alarme déclenchée » qui aura une mesure de confiance de 80%.

Par conséquence, nous pouvons constater que l'alarme est quatre fois plus probable qu'elle soit déclenchée ($S = 1$) que le contraire ($S = 0$).

Ainsi, l'évidence virtuelle apparaît en créant la variable G qui illustre le témoignage de Mme Gibbons comme indique la Figure 1. Dans ce cas, on peut considérer $P(g | s) = 0.4$ et $P(g | \text{not}(s)) = 0.1$, puis on insère l'observation $G = g$ dans le RB.

Contrairement à l'évidence virtuelle, qui peut être interprété comme une évidence avec incertitude, et représenté par un rapport de vraisemblance, « soft evidence » peut être interprété comme une évidence de l'incertitude et représenté par une distribution de probabilité d'une ou plusieurs variables.

B. Soft Evidence

Soft Evidence a été introduite par Voltora [12] qui l'a caractérisée par la distribution de probabilité $R(Y)$, $Y \subseteq X$ avec Y est un sous ensemble des variables X du RB. Ainsi, ce type d'évidence peut être vu de plusieurs cotés.

Par exemple, on ne peut pas observer l'état précis d'une variable mais plutôt on peut savoir la distribution de probabilité de ses états. De plus, parfois il est plus important de connaître la distribution de probabilité des états d'une variable que son état précis.

En outre, lorsque deux RBs interagissent entre eux, les informations échangées sont souvent sous la forme d'une distribution de probabilité de la variable partagée.

Concernant soft evidence, on constate une incertitude dans le choix de l'état observé X_i , mais on est certain de la distribution de probabilité. Autrement dit, la distribution de probabilité $R(X_i)$ est une observation certaine qui sera réservée pour la mise à jour de croyance.

Cependant, soft evidence sera traitée de la même façon que l'évidence classique. En effet, cette dernière qui impose $X_i = a$ est un cas particulier de soft evidence ($R(X_i = a) = 1$, $R(X_i = b) = 0$ pour tous états $b \neq a$).

Ainsi, le concept de soft evidence consiste à propager les valeurs quantifiées pour chaque état du nœud observé.

À ce sujet, Chan et Darwiche ont proposé dans [11] une nouvelle technique qui consiste à convertir soft evidence à l'évidence virtuelle.

Pour mieux expliquer ce concept, on peut illustrer un exemple (cf. Figure 2) détaillé dans [13].

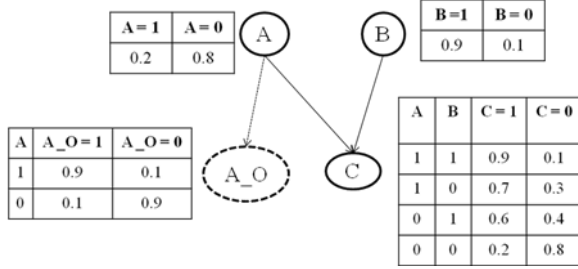


Figure 2. Exemple de RB avec Soft Evidence

Posons $Q(X)$ la distribution de valeurs quantifiées. Ainsi, si l'observation est effectuée sur le nœud A avec les pourcentages 90% pour l'état 1 et 10% pour l'état 2 alors les valeurs quantifiées seront comme suit :

$$\begin{aligned}
 Q(A=1) &= Q(A=1|A_o=1) \\
 &= [P(A_o=1|A=1) * P(A=1)] \\
 &/ [P(A_o=1|A=1) * P(A=1) + P(A_o=1|A=0) * P(A=0)] \\
 &= [0.9 * 0.2] / [0.9 * 0.2 + 0.1 * 0.8] \\
 &= 0.6923.
 \end{aligned}$$

$$\begin{aligned}
 Q(A=0) &= Q(A=0|A_o=1) \\
 &= [P(A_o=1|A=0) * P(A=0)] \\
 &/ [P(A_o=1|A=1) * P(A=1) + P(A_o=1|A=0) * P(A=0)] \\
 &= [0.1 * 0.8] / [0.9 * 0.2 + 0.1 * 0.8] \\
 &= 0.3077.
 \end{aligned}$$

Les évidences incertaines comportent deux catégories : évidence virtuelle et soft evidence. Mais sont dépourvus de moyen pour exprimer l'ambiguïté. Face à ce problème, nous avons proposé l'évidence floue dans les réseaux bayésiens.

IV. APPROCHE PROPOSEE POUR L'EVIDENCE FLOUE

A. Evidence Floue

Pour développer notre algorithme pour l'évidence floue dans les RBs, nous avons utilisé l'algorithme arbre de jonction combiné avec soft evidence modifié.

Cependant on peut résumer cet algorithme dans les deux étapes suivantes :

1. Pour chaque état i du nœud observé, on fait une inférence classique en observant l'état i .

2. Pour chaque nœud N du RB

Pour chaque état j de N

Etat $j = \sum$ valeur de l'état j issu de l'inférence i * degré d'appartenance de la valeur observée à l'état i

Fin Pour

Fin Pour

B. Solution Logicielle Proposée

Notre application, qui est développée en C#, est passée par trois étapes : tout d'abord, nous avons réalisé un éditeur graphique "Bayesian Editor" pour la représentation des réseaux bayésiens, puis, nous avons implémenté l'algorithme d'inférence JLO [14, 15] pour assurer la propagation de l'information dans le réseau, enfin, nous avons développé le module flou qui intègre la notion de l'évidence floue.

Afin d'étendre l'utilisation de notre éditeur au RB avec évidence floue, nous pouvons accéder au module de représentation des variables.

Cette interface inclut un tableau contenant les différentes données liées à la variable floue tel que son nom, sa fonction, l'intervalle de définition de cette fonction et la couleur qui sera choisie dans sa représentation graphique.

Dans ce tableau, nous pouvons ajouter, insérer ou supprimer des lignes à travers les boutons "Ajouter", "Insérer" et "Supprimer".

Cependant, les fonctions de la variable floue incluent plusieurs types de fonctions mathématiques notamment les nombres négatifs, les racines carrées, les puissances, sinus, cosinus, tangente, +, -, * et /.

Finalement, lors d'une évidence floue, nous pouvons saisir la valeur observée pour obtenir les degrés d'appartenance de chaque état de la variable floue dans le deuxième tableau.

V. EXEMPLE

Prenons un exemple détaillé dans [16], nous avons quatre variables : Neige (N), Encombrement (E), Tomber (T) et Arriver en retard (R) comme indique la Figure 3.

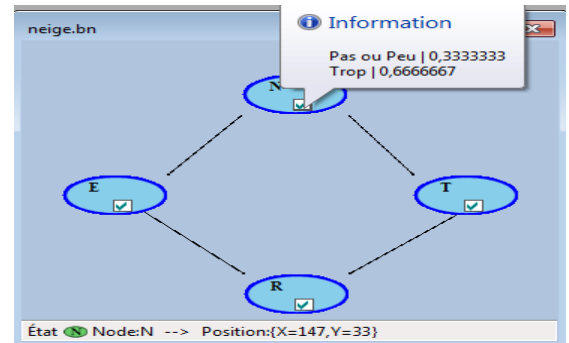


Figure 3. Probabilité à postériori

Ainsi, supposons que la variable « N » est une variable floue qui comporte deux sous-ensembles flous « Pas ou Peu de neige » et « Trop de neige » caractérisés par des fonctions bien déterminées qu'on peut saisir dans le tableau spécifique de la variable floue (cf. Figure 4).

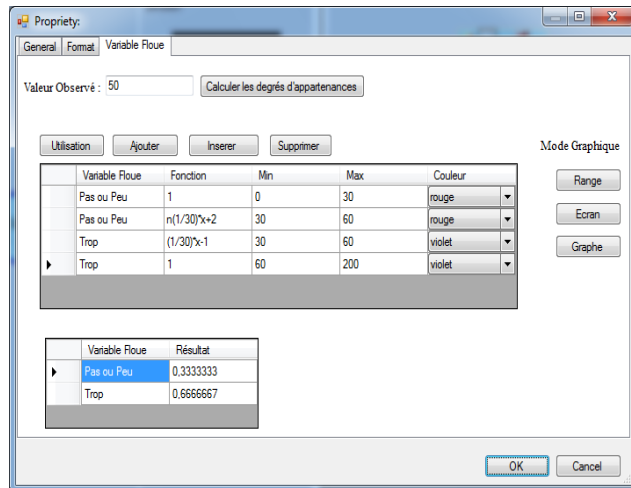


Figure 4. Représentation de la variable floue Neige

Supposons que la quantité de la neige observée est 50. Après réalisation de notre inférence, nous obtenons les probabilités *à posteriori* de chaque nœud (cf. Figure 3).

VI. CONCLUSION

Cet article présente une proposition d'intégration d'un nouveau type d'évidence dans les réseaux bayésiens celui de l'évidence floue.

Ainsi, nous avons pu en premier lieu modéliser quelques systèmes par réseaux bayésiens avec évidence floue, ce qui nous a permis de développer un module supplémentaire de représentation des variables floues à un éditeur graphique de représentation de RB en deuxième lieu, ainsi que de développer un algorithme exploitant la richesse de cette récente notion, puis tester et comparer l'algorithme que nous avons développé par un exemple illustratif et par rapport à d'autres travaux.

REFERENCES

- [1] P. Naïm, P. Willemin, P. Leray, O. Pourret, and A. Becker, "Réseaux bayésiens", 2004.
- [2] F. Jensen, "An Introduction to Bayesian Networks", Springer, New York, 1996.
- [3] O. Pourret, P. Naim and B. Marcot. Bayesian Networks: "A Practical Guide to Applications". Chichester, UK: Wiley. ISBN 978-0-470-06030-8, 2008.
- [4] L. Zadeh, "Fuzzy sets", Information and Control, Vol.8, 1965, 338-353.
- [5] J. Pearl, "Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference", Morgan Kaufmann, San Mateo, CA, 1988.
- [6] C. Huang and A. Darwiche. "Inference in belief networks: A procedural guide". International Journal of Approximate Reasoning, 15(3) : 225-263, 1996.
- [7] X. Li, "On the Use of Virtual Evidence in Conditional Random Fields", Empirical Methods in Natural Language Processing-EMNLP, pp. 1289-1297, 2009.

- [8] J. Bilmes. "On soft evidence in Bayesian networks". Technical Report UWEETR-2004-0016, University of Washington, Dept. of Electrical Engineering. 2004.
- [9] S. Reynolds and J. Bilmes. Part-of-speech tagging using virtual evidence and negative training. In Proceedings of HLT/EMNLP. 2005.
- [10] J. Pearl, Jeffery's rule, passage of experience, and neo-Bayesianism, in Knowledge Representation and Defeasible Reasoning, eds. H. E. Kyburg, Jr., R. P. Loui and G. N. Carlson (Kluwer Academic Publishers, Boston, 1990), pp. 245–265.
- [11] H. Chan and A. Darwiche. On the revision of probabilistic beliefs using uncertain evidence, Artificial Intelligence 163 (2005) 67–90.
- [12] M. Valtorta, Y. Kim and J. Vomlel, "Soft evidential update for probabilistic multiagent systems", International Journal of Approximate Reasoning 29(1) (2002) 71–106.
- [13] H. Tu, J. Allanach, S. Singh, K. R. Pattipati, and P. Willett. "The Adaptive Safety Analysis and Monitoring System". Electrical and Computer Engineering Department, University of Connecticut, 2004.
- [14] S. L. Lauritzen and D. J. Spiegelhalter. "Local computations with probabilities on graphical structures and their application to expert systems". International Journal of the Royal Statistical Society, pages 157.224, 1988.
- [15] F. Jensen, S. Lauritzen, and K. Olesen. "Bayesian updating in recursive graphical models by local computations". Computational Statistical Quarterly, 4 :269–282, 1990.
- [16] H. Tang, and S. Liu. Basic Theory of Fuzzy Bayesian Networks and Application in Machinery Fault Diagnosis. *Fourth International Conference On Fuzzy Systems and knowledge Discovery*. Haikou : IEEE, 2007.