De l'information à l'observation dans un réseau bayésien

Véronique Delcroix[1] **Ali Ben Mrad**[1,2] **Sylvain Piechowiak**[1]

[1]LAMIH UMR CNRS 8201, Université de Valenciennes et du Hainaut Cambrésis, 59313 Valenciennes FRANCE

[2] ENIS, CES, Sfax, University of Sfax, Tunisia

Veronique.delcroix@univ-valenciennes.fr

Résumé

Cet article est basé sur un état des lieux de la terminologie et des concepts des différents types d'observations dans les réseaux bayésiens. Il en résulte que ni la terminologie ni les définitions et concepts associés ne sont clairement établis et partagés, aussi bien dans la littérature que dans les logiciels sur les réseaux bayésiens. Nous proposons une présentation des concepts et du vocabulaire concernant les différents types d'observations dans les réseaux bayésiens. Nous identifions trois concepts d'observations incertaines : les observations de vraisemblance, les observations probabilistes fixes et les observations probabilistes non fixes. Chaque concept est défini, présenté et illustré à l'aide d'exemples.

Abstract

This paper is based on a review of the terminology and the associated concepts about the differents types of evidence in a Bayesian network. The result is that neither the terminology nor the definitions are clearly established, both in the literature and in Bayesian network software. We propose a presentation of the concepts and the terms about the different types of evidence in a Bayesian network. We identify three concepts of non deterministic evidence (or uncertain evidence) : likelihood evidence (or virtual evidence), fixed probabilistic evidence and not-fixed probabilistic evidence. Each concept is defined, explained and illustrated with several examples.

1 Introduction

Les réseaux bayésiens [24, 16, 3] sont des modèles graphiques probabilistes qui permettent de stocker de la connaissance sur un ensemble de variables et de mettre à jour l'état des croyances sur certaines variables cibles, en fonction des informations disponibles sur d'autres variables du modèle. La connaissance enregistrée dans un réseau bayésien inclut différentes formes d'incertitude : l'incertitude aléatoire qui représente la variabilité pour la population considérée et l'incertitude épistémique qui vient du manque de connaissance ou de données. L'état des croyances à un moment donné est exprimé sous la forme de distributions de probabilités a posteriori. Elles sont obtenues par la propagation des observations sur certaines variables du modèle grâce à des algorithmes d'inférence adaptés au type d'observation. Un réseau bayésien constitue une description d'un ensemble de connaissances, représenté par un graphe sur un ensemble de variables et par une distribution de probabilités P sur l'ensemble des variables du modèle. Plus précisément, étant donnés un graphe dirigé sans circuit $G = (\mathbf{X}, E)$ où $\mathbf{X} = X_1, ..., X_n$ est un ensemble de variables (les nœuds du graphe) et $\mathbf{E}$ un ensemble d'arcs, et P une distribution de probabilités sur l'ensemble des variables aléatoires $\mathbf{X}$ associées aux nœuds du graphe, on dit que (G, P) est un réseau bayésien (RB) si et seulement si :

$$P(X_1, ..., X_n) = \prod_{i=1}^{n} P(X_i|Pa(X_i)) \qquad (1)$$

où $Pa(Xi)$ est l'ensemble des parents dans le graphe G (nœuds reliés à X_i par des arcs d'extrémité X_i).

La définition d'un réseau bayésien permet de répondre à un ensemble de questions, représentées par des distributions de probabilité *a posteriori* sur des variables cibles sachant les valeurs des variables connues. Ainsi l'état des connaissances mises à jour est représenté par $P(. \mid e)$ où e représente les informations recueillies sur un sous-ensemble de variables. Ces informations constituent les données d'entrée du processus d'inférence, qui permet le calcul des distributions de probabilités *a posteriori*. Le plus souvent, ces informations correspondent à l'observation claire et directe d'un ensemble de variables du modèle. La valeur de ces variables étant connues avec certitude, ces observations se traduisent par l'instanciation des variables observées. Dans les cas où l'information reçue ne correspond pas à ce type de spécification, alors qu'elle est de nature à modifier l'état des croyances sur les variables du modèle, on parle d'observation incertaine. Deux types d'observations incertaines ont été identifiées dans le cadre plus général des différentes théories de l'incertain [14]

Cet article présente un état des lieux sur la terminologie des différents types d'observations dans les réseaux bayésiens, dans la littérature d'une part et dans les principaux logiciels concernant les RB d'autres part. Les constats sont les suivants :

1. le concept classique d'observation dans un réseau bayésien est largement partagé, malgré un ensemble de termes variés pour le désigner.
2. le concept d'observation de vraisemblance (likelihood evidence ou virtual evidence) proposé par Pearl [26] est clairement identifié et reconnu par la grande majorité des auteurs, même s'il est parfois mal nommé. La propagation de ce type d'observation est disponible dans de nombreux logiciels sur les réseaux bayésiens, avec toutefois des appellations variées et parfois source de confusion.
3. Une définition plus large du concept d'observation (finding) proposée dans [16] permet d'inclure le cas des observations négatives alors qu'il s'agit d'observations de vraisemblance.
4. un autre type d'observations appelé *soft evidence* en anglais mais non nommé en français est identifié dans la littérature, mais il y a parfois confusion dans les termes. Ce concept est globalement absent des logiciels sur les réseaux bayésiens.
5. un dernier concept d'observation correspond à l'information considérée dans la règle de Jeffrey pour réviser une distribution de probabilités jointe à partir d'une distribution de probabilités locale. Ce concept est mal identifié dans le cadre des réseaux bayésiens et rarement distingué du concept de soft evidence (sauf dans [2] où les trois types d'observations incertaines sont identifiés dans le même article.

Ces constats sont visibles aussi au niveau des logiciels sur les réseaux bayésiens où la terminologie des différents types d'observations incertaines est très variée et où les fonctionnalités pour propager ces observations sont peu présentes.

L'objectif de cet article est de clarifier les termes et les concepts concernant les différents types d'observations dans un réseau bayésien. Le tableau 1 présente une vue d'ensemble du vocabulaire proposés et des différents types d'observations dans un réseau bayésien. Nous identifions trois concepts d'observations incertaines : les observations de vraisemblance et les observations probabilistes fixes et non fixes.

La suite de cet article présente les définitions de ces différents types d'observations dans les RB puis dresse l'état des lieux de la terminologie sur les observations dans les RB, d'une part dans la littérature, d'autre part dans les logiciels sur les RB.

2 Terminologie proposée et définitions des observations

Cette partie présente et définit chaque type d'observation du tableau 1 avec plusieurs exemples de chaque type.

2.1 Les observations certaines

Il s'agit des observations couramment utilisées dans les réseaux bayésiens. Ce type d'observation est nommé information *déterministe* dans [24].

Definition 2.1 (Observation certaine). *Dans un réseau bayésien, une observation certaine e sur une variable X à valeur dans $\mathcal{D}_X$ est définie par un vecteur d'observation de taille $m = |\mathcal{D}_X|$ contenant un unique 1, à la position correspondant à l'état $x \in \mathcal{D}_X$ et des 0 aux positions des autres états de X. Cette observation représente l'instanciation de X avec la valeur x et est caractérisée par $P(X = x \mid e) = 1$.*

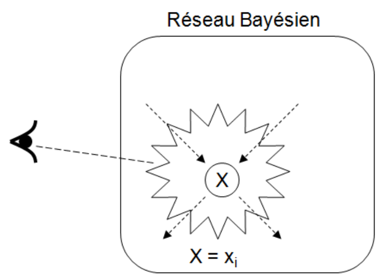

FIGURE 1 – Observation certaine.

La figure 1 illustre une observation certaine sur une variable X d'un réseau bayésien. La variable X est

Observation certaine	Observation incertaine		
	Observation de vraisemblance	Observation probabiliste non fixe	Observation probabiliste fixe
instanciation	observation avec incertitude (capteur imprécis) + connaissance du lien entre la valeur réelle et la valeur observée	observation hors du réseau bayésien + connaissance de son influence sur les variables du Réseau bayésien.	(1) observation + transformation (observation d'une variable continue + discrétisation floue) (2) regroupement d'observations sur la même variable (sous-population) (3) observations hors du modèle + combinaison des influences sur une variable du modèle (modèle AEBN [5, 22])
pas d'incertitude	l'incertitude porte sur l'observation	l'incertitude porte sur la connaissance ajoutée	l'incertitude porte sur la transformation de l'information initiale en une observation sur une variable du modèle

TABLE 1 – Synthèse des différent types d'observations dans un réseau bayésien

observée directement et sans incertitude. Cette observation permet d'affirmer que la variable X est dans l'état x_i.

Exemple 2.1 (Observations certaine)**.** *Considérons un réseau bayésien contenant les variables A et S représentant l'âge et le sexe d'une personne, avec les ensembles de définition $\mathcal{D}_A = \{[0,11],[12,25],[26,65],[66,130]\}$ et $\mathcal{D}_S = \{M,F\}$. Le cas d'une jeune fille de 20 ans se traduit par les deux observations $A = [12,25]$ et $S = F$.*

La suite de cette partie présente les trois types d'observations incertaines.

2.2 Les observations de vraisemblance

Les observations de vraisemblance (en anglais *likelihood evidence* ou *virtual evidence*) concernent les cas où l'observation d'une variable est incertaine. Il existe une incertitude sur l'observation elle même, par exemple parce l'information est fournie par une source qui n'est pas parfaitement fiable, ou précise. Pour spécifier cette observation, il est nécessaire d'avoir une connaissance, au moins partielle, du lien entre les différentes valeurs réelles possibles de la variable et la valeur observée.

Definition 2.2 (Observations de vraisemblance)**.** *Une observation de vraisemblance sur une variable X d'un réseau bayésien est spécifiée par un rapport de vraisemblance*[1] *$L(X) = (L(X = x_1) : \ldots : L(X = x_m))$, où les $L(X = x_i)$ sont des quantités relatives et le rapport de vraisemblance est défini par*

$$L(X) = (P(Obs \mid x_1) : \ldots : P(Obs \mid x_n))$$

où $P(Obs \mid x_i)$ est interprété comme la probabilité de l'évènement observé sachant que X est dans l'état x_i.

Une observation de vraisemblance sur X spécifiée par un vecteur de 0s et de 1s est parfois appelée *observation négative* pour souligner que les états de X correspondant aux 0s sont des états impossibles. Une telle observation pourrait aussi être appelée *observation disjonctive*, pour souligner que la variable X se trouve dans l'un ou l'autre des états correspondant aux 1s.

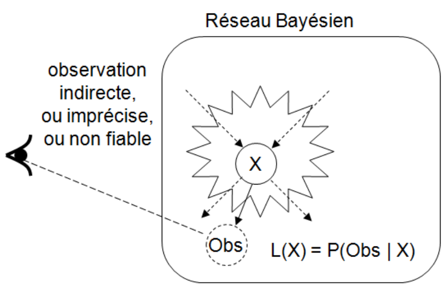

FIGURE 2 – Observation de vraissemblance sur X.

La figure 2 illustre un exemple type d'observation de vraisemblance sur une variable X d'un réseau bayésien. La variable X n'est pas observée directement, avec certitude. L'observation permet néanmoins d'apprendre quelque chose sur la valeur réelle de X. Cette information sur X apportée par l'observation est spécifiée par un rapport de vraisemblance de l'observation par rapport à chaque valeur de X.

Exemple 2.2 (Observations de vraisemblance)**.** *Considérons un réseau bayésien contenant une variable D représentant la valeur d'un dé à six faces.*

1. Les valeurs d'un rapport de vraisemblance sont séparées par des deux-points, qui rappellent que seuls comptent les rapports entre les valeurs. De ce fait, il n'y a pas de condition de normalisation pour un rapport de vraisemblance.

L'observateur est un jeune enfant qui confond encore certains chiffres, et il déclare "c'est un six". Connaissant le niveau d'acquisition des chiffres de l'enfant, on estime qu'il y a deux fois plus de chances que l'enfant déclare avoir vu un six si le dé est effectivement tombé sur la face numéro six que s'il est tombé sur la face trois, quatre ou cinq ; et on estime que l'enfant n'affirmera pas cela si le dé est tombé sur la face un ou deux car il connait ces chiffres. Cette information (déclaration de l'enfant + notre connaissance de l'enfant) peut être spécifiée par le rapport de vraisemblance $L(X) = (0 : 0 : 1 : 1 : 1 : 2)$, *qui indique entre autre que* $P(Obs \mid D = six) = 2 \times P(Obs \mid D = quatre)$.

La méthode de Pearl pour propager une observation de vraisemblance sur X consiste à ajouter virtuellement et temporairement une variable *Obs* à deux états (*obs* et *non obs*) dont l'unique parent est X et dont la table de probabilité conditionnelle de la nouvelle variable sachant X est définie par le rapport de vraisemblance (à un coefficient près [2]). Il suffit ensuite de propager l'observation certaine $Obs = obs$.

Lorsqu'une seule valeur est possible, il s'agit d'une observation certaine.

Exemple 2.3 (observation négative ou observation disjonctive)**.** *Considérons de nouveau l'exemple du dé à six faces. L'information "le dé est tombé sur un nombre pair" se traduit par le vecteur d'observation* $L(D) = (0 : 1 : 0 : 1 : 0 : 1)$ *qui traduit le fait que le dé est tombé sur la face deux, quatre ou six, et que les faces un, trois et cinq sont exclues.*

2.3 Les observations probabilistes fixes et non fixes

Le concept d'observation probabiliste (fixe et non fixes) est parfaitement décrit dans [14] : l'information est une description partielle d'une mesure de probabilité et constitue une contrainte sur l'état cognitif final. L'information vient donc corriger (remplacer) l'état cognitif initial.

Definition 2.3 (Observation probabiliste (fixe et non fixe))**.** *Une* observation probabiliste *sur une variable* $X \in \mathbf{X}$ *dans un réseau bayésien est spécifiée par une distribution de probabilité* $R(X)$ *qui définit une contrainte sur la variable* X *après que cette information ait été propagée ; cette distribution de probabilité décrit l'état des croyances sur la variable* X *"toutes choses étant considérées". Une observation probabiliste est* fixe *(non fixe) lorsque la distribution de probabilités* $R(X)$ *ne peut pas (peut) être modifiée par la propagation d'autres observations.*

2. ce coefficient n'a aucun effet sur la propagation de *Obs*.

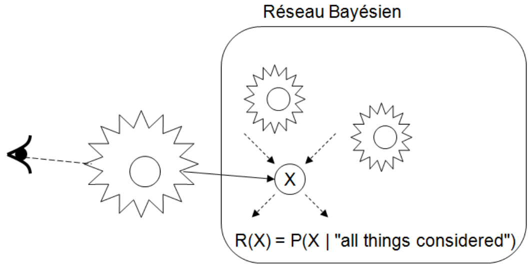

FIGURE 3 – Observation probabiliste non fixe sur X.

Les observations probabilistes non fixes correspondent au type d'information utilisées dans la règle de Jeffrey [15] pour la révision d'une distribution de probabilité jointe par une distribution locale de probabilité sur un sous-ensemble de variables. La figure 3 illustre un exemple d'observation probabiliste non fixe sur une variable X d'un réseau bayésien. La nouvelle information inclut l'observation d'une variable qui ne fait pas partie du réseau bayésien, ainsi que la connaissance de la façon dont cette observation influe sur X. L'observation probabiliste non fixe sur X est spécifiée par une distribution de probabilité $R(X)$ qui intègre l'influence de toutes les observations disponibles. Cette distribution de probabilité locale remplace la distribution de probabilité a priori $P(X)$ aussi longtemps que l'information est valide. La propagation de cette information par la règle de Jeffrey aboutit à une distribution de probabilité Q sur l'ensemble des variables du modèle, telle que $Q(X) = R(X)$ et les distributions de probabilité conditionnelles des autres variables sachant X restent inchangées. Cette information n'empêche pas que d'autres informations puissent modifier nos croyances sur X.

La propagation d'une observation de vraisemblance avec la méthode de Pearl fournit les même résultats que la propagation d'une distribution locale de probabilités par la règle de Jeffrey, après avoir converti l'observation probabiliste non fixe en observation de vraisemblance [10].

Exemple 2.4 (Observations probabiliste non fixe)**.** *(suite de l'exemple du dé à six faces). Comment traduire l'information "il s'agit d'un dé truqué" si le modèle ne contient pas de variable pour représenter la nature du dé ? Cette information concerne l'observation d'une variable extérieure au modèle, mais qui modifie nos croyances sur les valeurs du dé. Cette information peut être représentée par une observation probabiliste non fixe sur la variable* D, *à condition d'y associer un élément de connaissance supplémentaire, comme le fait que le dé truqué favorise l'apparition du six, qui peut se traduire par une distribution de probabilité* $R(D) = P(D \mid déTruqué = vrai$. *L'information que le* *dé est truqué peut être combinée avec une autre information sur X (observation de vraisemblance ou autre observation probabiliste non fixe).*

Considérons maintenant une *observation probabiliste fixe* : elle représente une information qui ne peut pas être remise en cause. De ce fait, elle se comporte en tous points comme une observation certaine. Une observation probabiliste fixe sur une variable X (1) définit une contrainte sur la distribution de probabilité de X qui ne peut être modifiée par aucune autre information, (2) ne peut être combinée avec une autre information sur la même variable, (3) remplace tout *a priori* sur X et enfin, (4) la propagation de plusieurs observations probabilistes fixes sur différentes variables du modèle aboutit au même résultat quel que soit l'ordre de leur propagation : il y a commutation. Ce dernier point constitue une différence essentielle entre les observations probabilistes fixes et non fixes. C'est le caractère "certain" des observations probabilistes fixes qui les différencie essentiellement des observations probabilistes non fixes. Au contraire, une observation probabiliste non fixe sur une variable peut être modifiée par l'arrivée d'autres informations.

Les figures 4, 5 et 6 illustrent trois exemples d'observation probabiliste fixe sur une variable X.

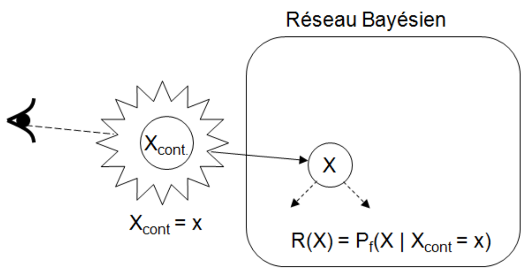

FIGURE 4 – Observation probabiliste fixe sur X (exemple 2.5) : l'observation d'une variable continue est transformée à l'aide d'une discrétisation floue.

Dans l'exemple2.5 (figure 4), l'information vient de l'observation d'une variable continue ; la variable discrète X du réseau bayésien est associée à la variable continue par une partition floue. L'observation probabiliste fixe sur X est spécifiée par une distribution de probabilités $R(X)$ qui projette la valeur observée sur la partition floue.

Exemple 2.5 (Observations probabiliste fixe : discrétisation floue)**.** *Considérons un réseau bayésien contenant une variable A (âge), à valeur dans $\mathcal{D}_A = \{enfant, adulte, personneAgee\}$. L'information "le sujet a 15 ans" peut se traduire par une observation probabiliste fixe sur la variable A de notre réseau bayésien, si on y ajoute un élément de connaissance concernant la fonction floue qui associe à chaque valeur réelle de l'âge un degré d'appartenance aux intervalles flous de $\mathcal{D}_A$. On a par exemple $R(A) = (0.7, 0.3, 0)$. Cette observation probabiliste sur A est fixe car aucune autre observation ne peut modifier cette distribution.*

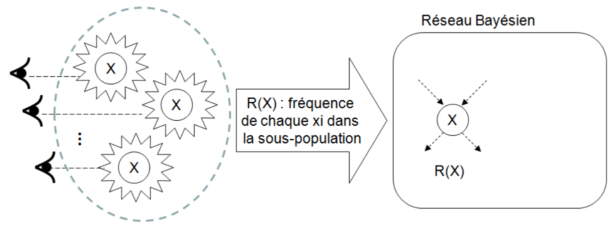

FIGURE 5 – Observation probabiliste fixe sur X (exemple 2.6) : observation d'une sous-population.

Dans l'exemple 2.6 (figure 5), l'information vient d'un ensemble d'observations de X, par exemple issu de l'observation de chaque individu d'une sous-population. L'observation probabiliste fixe sur X est spécifiée par une distribution de probabilité $R(X)$ qui représente la fréquence d'apparition de chaque valeur dans la sous-population observée.

Exemple 2.6 (Observations probabiliste fixe : observation d'une sous-population)**.** *Considérons le réseau bayésien Asia [23], qui contient huit nœuds binaires, dont une nœud racine S (Smoking) et un nœud feuille D (Difficultés respiratoires). L'information concerne les travailleurs de telle usine et indique que la moitié d'entre eux souffre de difficultés respiratoires, ce qui peut se traduire par une observation probabiliste fixe sur D : $R(D) = (0.5, 0.5)$. Une deuxième information indique que seul un dixième d'entre eux sont des fumeurs : $R(S) = (0.1, 0.9)$. La première information doit être préservée par la propagation de la deuxième information car chacune des observations probabilistes fixes doit restée inchangée. Si l'ensemble des couples d'observations est disponible pour chaque individu de la sous-population, on peut considérer une seule observation probabiliste fixe $R(D, S)$ [30].*

Dans l'exemple 2.7 (figure 6), la variable X est commune à deux réseaux bayésiens : le premier est un "agent" expert de la variable X, alors que le second "agent", lui aussi doté d'un réseau bayésien local est en quête d'une évaluation de la variable X. Le premier réseau bayésien évalue la valeur de X grâce à un ensemble d'observations locales et envoie son évaluation $R(X)$ au second agent. Ce dernier met à jour son réseau bayésien local à partir de l'observation probabiliste fixe sur X spécifiée par $R(X)$. Il s'agit du modèle Agent Encapsulated Bayesian Network (AEBN) [5, 22].

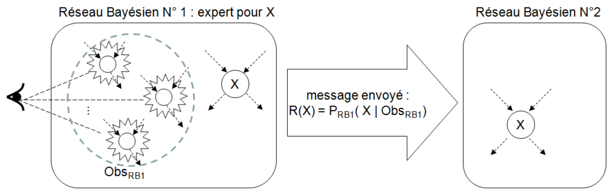

FIGURE 6 – Observation probabiliste fixe sur X (exemple 2.7) : un "agent" expert sur X envoie son évaluation de X à un autre agent qui la lui demande.

Exemple 2.7 (Observations probabiliste fixe : AEBN)**.** *Considérons deux agents, chacun doté d'un réseau bayésien local comme modèle de connaissance et de raisonnement. Le premier agent est spécialiste de l'évaluation des capacités motrices d'une personne handicapée. Le second agent a pour objectif de fournir des recommandations en vue du choix d'un fauteuil roulant [13]. Il fait appel au premier pour obtenir une évaluation de CM la capacité de marche de la personne qui est une variable partagée par ces deux agents. Le premier agent actualise les distribution de probabilité de la variable CM en fonction des informations recueillies sur la personne, et la transmet au second agent qui la reçoit, et la propage "chez lui" sans la modifier.*

Le tableau 2 présente une synthèse des caractéristiques des différents types d'observations dans un réseau bayésien présentés dans cette partie.

3 Etat des lieux sur les observations dans les RB

Nous présentons dans cette partie l'état des lieux sur les termes et les concepts concernant les observations dans les réseaux bayésiens, d'une part dans la littérature, et d'autre part dans les logiciels sur les réseaux bayésiens.

3.1 Les termes employés pour désigner une observation certaine

Le terme observation [3] est suffisant lorsqu'il n'est question que d'observations certaines. Les termes *observation déterministe* sont utilisés dans [24, 6]. Pour être plus précis, les termes les plus fréquemment utilisés en anglais sont *hard finding* et *hard evidence*, qui n'ont pas d'équivalent exacts en français. On trouve aussi parfois les termes *specific finding*, *regular finding* et *positive finding* ou le terme evidence avec ces qualificatifs.

Une définition plus générale du terme anglais *finding* proposée dans [16] qui correspond à une représentation par un vecteur pouvant contenir plusieurs uns pour les valeurs possibles, et des zéros pour les valeurs impossibles. Nous ne retenons pas cette généralisation car elle inclut dans le terme *finding* une acceptation particulière des observations de vraisemblance.

La suite de cette revue de la littérature concerne plus spécifiquement les observations incertaines.

3.2 Usage des termes "observations incertaines"

Les termes *observations incertaines* sont généralement employés pour désigner différents types d'observations non classiques, ou l'un de ces types. Dans [14] les termes *uncertain input* désignent les deux grands deux types d'observations incertaines dans le cadre plus général des différentes théories de l'incertain. Dans [2], les termes *uncertain evidence* désignent explicitement les trois types d'observations incertaines dans un réseau bayésien. Dans [1], le terme *uncertain information* désigne un type précis d'observation incertaine (*virtual evidence*). Dans [10, 12], les termes *uncertain evidence* regroupent deux types d'observations incertaines qui sont les observations de vraisemblances et les observations probabilistes non fixes, alors que dans [31, 25, 21, 27] ces même termes regroupent les observations de vraisemblance et les observations probabilistes fixes.

Un dernier mot sur la place de l'incertitude par rapport à l'observation : dans [27], les observations de vraisemblance sont désignées comme *observations avec incertitude* (evidence **with** uncertainty), alors que les observations probabilistes fixes *soft evidence* sont présentées comme des *observations de l'incertidude* (evidence **of** uncertainty).

3.3 A propos des termes anglais soft evidence

Le terme anglais *soft evidence* a été introduit dans [30], bien que le concept ait été identifié dans [5]. Cependant, en dehors du groupe des auteurs associés à Valtorta, peu d'articles utilisent ce terme pour désigner des observations probabilistes fixes [29, 19]. Jusqu'à présent, il n'existe pas de termes en français pour désigner ce concept, et c'est pourquoi nous proposons le terme d'observation probabiliste fixe.

En revanche, de nombreux articles utilisent les termes *soft evidence* pour désigner un autre concept : dans [18, 20, 4, 11, 9, 8], ils sont employés pour désigner des observations de vraisemblance [4]. Dans [7],

3. en anglais *observations*, *findings* ou *evidence* : nom commun indénombrable (sans pluriel)

4. L'usage des termes *soft evidence* est abandonné dans [10]

	Observation certaine	Observation incertaine		
		Observation de vraisemblance	Observation probabiliste non fixe	Observation probabiliste fixe
L'observation définit une contrainte sur la distribution de probabilité *a posteriori*. Elle est spécifiée "toutes choses considérées"	Oui	Non	Oui	Oui
Plusieurs observations peuvent être combinées sur la **même** variable	Non	Oui	Oui	Non
La distribution de probabilité *a posteriori* peut être modifiée par la propagation d'observations sur d'**autres** variables	Non	Oui	Oui	Non
La distribution de probabilité *a priori* sur la variable est utilisée lors de la propagation	Non	Oui	Non	Non
La propagation de plusieurs observations commute	Oui	Oui	Non	Oui

Table 2 – Synthèse des propriétés des différents types d'observations dans un réseau bayésien.

ils sont utilisés pour signifier qu'une variable ne peut prendre une certaine valeur (interdite), ce qui est habituellement désigné comme étant une *observation négative*.

Cette confusion sur le termes *soft evidence* est présente aussi dans les logiciels sur les réseaux bayésiens. Certains logiciels sur les réseaux bayésiens ne disposent d'aucune fonctionnalité concernant la propagation d'observation incertaine (par exemple Elvira [5], Analytica [6], Samlam [7]). Parmi les logiciels qui implémentent la méthode de Pearl pour la propagation des observations de vraisemblance, trois d'entre eux utilisent les termes *soft evidence* pour désigner des observations de vraisemblance (voir table 3).

Logiciel	version	termes utilisés pour désigner les observations de vraisemblance
BayesiaLab [8]	5.3	likelihoods
Netica [9]	5.12	likelihood finding [10]
Genie [11]	2.0	virtual evidence
Hugin [12]	8.0	likelihood evidence
BNT [13]	1.0.7	soft evidence
Infer.NET [14]	2.5	virtual evidence
Bayes Server [15]	5.5	soft evidence
AgenaRisk [16]	6.1	soft evidence

Table 3 – Terminologie utilisée dans les logiciels sur les réseaux bayésiens pour désigner les observations de vraisemblance (avril 2014).

La propagation des observations probabilistes (fixes ou non fixes) est quasiment absente des logiciels sur les réseaux bayésiens. La table 4 montre les termes employés pour désigner les observations probabilistes fixes, dans les deux seuls logiciels disponibles qui permettent la propagation des observations probabilistes fixes et/ou non fixes.

Logiciel (version)	Termes utilisés pour les observations probabilistes	
	non fixes	fixes
BayesiaLab [17] (5.3)	probability distribution (not fixed)	probability distribution (fixed)
Netica [18] (5.12)	calibration	

Table 4 – Terminologie utilisée dans les logiciels sur les réseaux bayésiens pour désigner les observations probabilistes fixes et non fixes (avril 2014).

5. http ://leo.ugr.es/elvira/
6. http ://www.bayesia.com
7. http ://reasoning.cs.ucla.edu/samiam
8. http ://www.bayesia.com
9. http ://www.norsys.com/
10. the term"soft evidence" was used in previous versions of Netica and is still used in the help
11. http ://genie.sis.pitt.edu
12. http ://www.hugin.com. BC-Hugin is an extension of Hugin that has been proposed in [17] to implement the Big Clique algorithm to propagate probabilistic evidence in a BN. However, this extension is no longer readily available.
13. http ://www.cs.ubc.ca/ murphyk/Software/BNT/bnt.html
14. http ://research.microsoft.com/en-us/um/cambridge/projects/infernet/default.aspx
15. http ://www.bayesserver.com/
16. http ://www.agenarisk.com/
17. http ://www.bayesia.com
18. http ://www.norsys.com/

Aucun des deux concepts d'observation probabiliste (fixe et non fixe) n'est présent dans les logiciels sur les réseaux bayésiens actuels, à de rares exceptions, ce qui est un indicateur de la méconnaissance générale de ces concepts. En conclusion, les termes *soft evidence* portent à confusion du fait qu'ils sont utilisés pour désigner différents types d'observations incertaines aussi bien dans la littérature que dans les logiciels et plus généralement, les concepts d'observations probabilistes restent peu connus dans la communauté des chercheurs et des utilisateurs des réseaux bayésiens.

3.4 Confusion entre les observations probabilistes fixes et non fixes dans la littérature et discussion sur la commutation

La notion d'observation probabiliste non fixe vient du cadre probabiliste et la comparaison avec avec la notion d'observation dans un réseau bayésien est bien étudié dans [10]. Les auteurs étudient et comparent la spécification et la propagation des observations de vraisemblance et des observations probabilistes non fixes, sans proposer de terme spécifique pour désigner ce deuxième type d'observation incertaine. Par ailleurs, le concept d'observation probabiliste fixe n'y est pas mentionné.

Le groupe des auteurs associés à Valtorta a beaucoup travaillé notamment sur le concept d'observations probabilistes fixes (soft evidence), et en particulier sur les algorithmes de propagation [30, 31, 17, 25, 21, 22, 27, 28]. Dans ces articles, la distinction avec les observations probabilistes non fixes n'est jamais clairement établie, bien que la règle de Jeffrey, qui permet de propager les observations probabilistes non fixes y soit présentée.

Cette confusion (ou non distinction) entre les observations probabilistes fixes et non fixes conduit par ailleurs les auteurs de [27] à faire état d'une divergence de point de vue concernant la question de la commutation suite à la propagation de plusieurs observations probabilistes (soft evidence dans [27]) dans un réseau bayésien. Dans [10], les auteurs examinent la question de la commutation après la propagation de plusieurs observations incertaines dans un réseau bayésien. Avec la méthode de Pearl, il y a commutation : plusieurs observations de vraisemblance peuvent être propagées successivement sans que l'ordre de leur propagation modifie le résultat sur l'état final des croyances (la distribution de probabilité a posteriori). En revanche, ce n'est pas le cas avec la règle de Jeffrey : lorsqu'il y a plusieurs observations spécifiées par une distribution locale de probabilité, la conversion de chacune en un rapport de vraisemblance et la propagation de chacune par la méthode de Pearl conduit a des résultats différents suivant l'ordre des observations. De plus, les distributions locales de probabilités utilisées pour spécifier chaque observations ne sont pas toutes préservées par les propagations multiples. Certains auteurs affirment que la propagation de plusieurs observations spécifiées "toutes choses étant considérées" (comme c'est le cas pour la règle de Jeffrey) ne doit pas être commutative [10] ; alors que d'autres expliquent que lorsque plusieurs observations probabilistes représentent chacune une véritable observation de la distribution de certains événements, alors chacune de ces distributions locales de probabilité doit être préservées dans la distribution de probabilité *a posteriori* après avoir pris en compte chacune de ces observations, et ce quel que soit leur ordre [27]. Dans le premier cas, une information sur une variable spécifiée par une distribution locale de probabilité est susceptible d'être modifiée par de nouvelles informations, alors que dans le second cas, l'information probabiliste doit se comporter comme une observation certaine et ne peut pas être modifiée par de nouvelles informations. Cette revue de la littérature conduit au constat que plusieurs types d'observations incertaines ont été identifié et définis, mais que ni la terminologie ni les concepts ne sont complètement et clairement défini dans la communauté des chercheurs et celle des utilisateurs des réseaux bayésiens. C'est ce constat qui a motivé la synthèse et les clarifications présentées dans la première partie de cet article.

4 Conclusion

Cet article propose une présentation des différents types d'observations dans les réseaux bayésiens, comprenant la terminologie, les définitions et des exemples de chaque type d'observation. L'identification de trois types différents d'observation incertaines n'est mentionné dans la littérature que dans un seul article à notre connaissance [2], qui ne détaille pas cette question. Bien que chaque type d'observation ait déjà été étudié, au moins partiellement, aucune synthèse n'existe à ce sujet. Plus précisément, c'est la distinction entre observation probabiliste fixe et non fixe qui est totalement méconnue [19] et plus généralement les deux notions d'observations probabilistes (fixes et non fixes) restent très peu connues des communautés des chercheurs et des utilisateurs des réseaux bayésiens.

Pour chaque type d'observation, nous avons présenté à l'aide de schémas la position de l'observation par rapport aux variables du réseau bayésien ; Ces schémas ne prétendent pas à l'exhaustivité, mais ils

19. à l'exception du logiciel Bayesia lab qui identifie chacun de ces concepts, sans en faire de présentation détaillée

permettent de mettre en valeur un aspect intéressant des observations incertaines : ces différents types d'observations permettent de prendre en compte des informations sur des variables non représentées dans le modèle, avec différentes configurations : soit une nouvelle variable est prise en compte en dehors du réseau bayésien (figure 3) soit une même variable est dupliquée : (1) une variable continue hors du modèle est associée à une variable discrétisée lui correspondant dans le modèle (figure 4) ; (2) une variable représentant une grandeur réelle dans le modèle est associée à une variable représentant l'observation (voir figure 2) ; (3) une variable est commune à deux réseaux bayésiens : un agent expert de cette variable envoie son évaluation de la variable à un autre agent qui la lui demande (figure 5) ; (4) dans le dernier cas (figure 6), la variable hors du modèle représente une grandeur pour un individu (un cas), alors que la variable dans le modèle est utilisée pour représenter cette même grandeur sur une sous-population. L'observation probabiliste fixe représente la répartition statistique de la variable dans la sous-population considérée.

Dans l'utilisation la plus classique d'un réseau bayésien, on exploite des informations qui sont des observations de variables appartenant au modèle. Dans les cas ci-dessus — à l'exception de la figure 6 qui combine un ensemble d'observations en une seule — la nouvelle information pourrait être spécifiée comme une observation certaine, à condition de remplacer le modèle par un autre plus complet, mais aussi plus complexe, de taille plus grande, etc. Il apparait donc que les observations incertaines permettent d'utiliser un réseau bayésien un peu au delà du cadre pour lequel il a été conçu, ce qui nous semble constituer une capacité intéressante.

A ce jour, plusieurs logiciels sur les réseaux bayésiens permettent la propagation des observations de vraisemblance, mais de diverses améliorations peuvent être apportées : aide à l'utilisateur de réseau bayésien pour la spécification ces observations ou encore possibilité de combiner plusieurs observations de vraisemblance sur une même variables (différents observateurs simultanés). En revanche, très peu de logiciel proposent la propagation d'observation probabilistes non fixes et un seul permet la propagation d'observation probabiliste fixe [20]. Il nous semblerait intéressant que les logiciels sur les réseaux bayésiens proposent des fonctionnalités correspondant aux trois types d'utilisation des observations probabilistes fixes décrites dans cet article, ainsi que la possibilité de spécifier et de propager des observations probabilistes non fixes.

20. Bayesialab

Remerciements

Nous remercions pour le soutien qu'ils ont apporté à ce travail le Ministère de l'Éducation Nationale, de la Recherche et de la Technologie, la Région Nord-Pas de Calais et le Centre National de la Recherche Scientifique. Nous tenons aussi à remercier Philip Leicester pour ses remarques sur une version longue en anglais de cet article.

Références

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

[2] Salem Benferhat and Karim Tabia. Inference in possibilistic network classifiers under uncertain observations. *Annals of Mathematics and Artificial Intelligence*, 64(2-3) :269–309, 2012.

[3] Pierre Bessière, Emmanuel Mazer, Juan Manuel Ahuactzin-Larios, and Kamel Mekhnacha. *Bayesian Programming*. CRC Press, December 2013.

[4] Jeff Bilmes. On soft evidence in Bayesian networks. Technical Report UWEETR-2004-00016, Department of Electrical Engineering University of Washington, Seattle, 2004.

[5] Mark Bloemeke. *Agent encapsulated Bayesian networks*. Ph.d. thesis, Department of Computer Science, University of South Carolina, 1998.

[6] Remco R. Bouckaert, Enrique Castillo, and José-Manuel Gutiérrez. A modified simulation scheme for inference in Bayesian networks. *International Journal of Approximate Reasoning*, 14(1) :55–80, 1996.

[7] C. J. Butz and F. Fang. Incorporating evidence in Bayesian networks with the select operator. In *In Proceedings of the 18 th Canadian Conference on Artificial Intelligence, Springer-Verlag, Victoria British-Columbia*, pages 297–301, 2005.

[8] Hei Chan. *Sensitivity Analysis of Probabilistic Graphical Models*. Ph.d. thesis, University of California, Los Angeles, 2005.

[9] Hei Chan and Adnan Darwiche. Sensitivity analysis in Bayesian networks : From single to multiple parameters. In *UAI*, pages 67–75, 2004.

[10] Hei Chan and Adnan Darwiche. On the revision of probabilistic beliefs using uncertain evidence. *Artificial Intelligence*, 163(1) :67–90, 2005.

[11] B. D'Ambrosio, M. Takikawa, and Upper D. Representation for dynamic situation modeling. Technical Report Technical report, Information Extraction and Transport, Inc., 2000.

[12] Adnan Darwiche. *Modeling and Reasoning with Bayesian Networks.* Cambridge University Press, 2009.

[13] Veronique Delcroix, Karima Sedki, and François Xavier Lepoutre. A Bayesian network for recurrent multi-criteria and multi-attribute decision problems : Choosing a manual wheelchair. *Expert Systems with Applications*, 40(7) :2541–2551, 2013.

[14] Didier Dubois, Serafin Moral, and Henri Prade. Belief change rules in ordinal and numerical uncertainty theories. In D.M. Gabbay and Ph. Smets, editors, *Belief Change, (D. Dubois, H. Prade, eds.), Vol. 3 of the Handbook of Defeasible Reasoning and Uncertainty Management Systems*, pages 311–392. Kluwer Academic Publishers, Dordrecht, 1998.

[15] Richard C. Jeffrey. *The Logic of Decision.* (2nd edition) University of Chicago Press, 1990. 246 pages.

[16] Finn Verner Jensen and Thomas D. Nielsen. *Bayesian Networks and Decision Graphs.* Springer Publishing Company, Incorporated, 2nd edition, 2007.

[17] Young-Gyun Kim, Marco Valtorta, and Jirí Vomlel. A prototypical system for soft evidential update. *Applied Intelligence*, 21(1) :81–97, 2004.

[18] U.B. Kjaerulff and A.L. Madsen. *Bayesian Networks and Influence Diagrams : A Guide to Construction and Analysis*, volume 22 of *Information science and statistics.* Springer, 2 edition, 2013.

[19] T. Koski and J. Noble. *Bayesian Networks : An Introduction.* Wiley Series in Probability and Statistics. Wiley, 2009.

[20] Mark L. Krieg. A tutorial on Bayesian belief networks. Technical Report DSTO-TN-0403, Surveillance Systems Division, Electronics and Surveillance Research Laboratory, Defense science and technology organisation, Edinburgh, South Australia, Australia, 2001.

[21] Scott Langevin and Marco Valtorta. Performance evaluation of algorithms for soft evidential update in Bayesian networks : First results. In *SUM*, pages 284–297, 2008.

[22] Scott Langevin, Marco Valtorta, and Mark Bloemeke. Agent-encapsulated Bayesian networks and the rumor problem. In *AAMAS '10 Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems*, volume 1, pages 1553–1554, 2010.

[23] Steffen L. Lauritzen and David J. Spiegelhalter. Local computations with probabilities on graphical structures and their application to expert systems. *Journal of the Royal Statistical Society, Series B*, 50 :157–224, 1988.

[24] Patrick Naïm, Pierre-Henri Wuillemin, Philippe Leray, Olivier Pourret, and Anna Becker. *Réseaux bayésiens.* Eyrolles, 3 edition, 2007.

[25] Rong Pan, Yun Peng, and Zhongli Ding. Belief update in Bayesian networks using uncertain evidence. In *ICTAI*, pages 441–444, 2006.

[26] Judea Pearl. *Probabilistic reasoning in intelligent systems : networks of plausible inference.* Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1988.

[27] Y. Peng, S. Zhang, and R. Pan. Bayesian network reasoning with uncertain evidences. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 18(5) :539–564, 2010.

[28] Yun Peng, Zhongli Ding, Shenyong Zhang, and Rong Pan. Bayesian network revision with probabilistic constraints. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 20(3) :317–337, 2012.

[29] E. Di Tomaso and J. F. Baldwin. An approach to hybrid probabilistic models. *International Journal of Approximate Reasoning*, 47(2) :202–218, 2008.

[30] Marco Valtorta, Young-Gyun Kim, and Jirí Vomlel. Soft evidential update for probabilistic multiagent systems. *International Journal of Approximate Reasoning*, 29(1) :71–106, 2002.

[31] Jirí Vomlel. Probabilistic reasoning with uncertain evidence. *Neural Network World, International Journal on Neural and Mass-Parallel Computing and Information Systems*, 14(5) :453–465, 2004.