

Réseaux causaux possibilistes pour le traitement des interventions

Salem Benferhat¹, Salma Smaoui¹

CRIL, Université d'Artois,
Rue Jean Souvraz SP 18, 62300 Lens, France
{benferhat, smaoui}@cril.univ-artois.fr

Résumé : Ce article présente deux principales contributions pour le développement des réseaux causaux possibilistes. La première concerne la représentation des interventions dans les réseaux possibilistes. Nous proposons une contrepartie de l'opérateur "DO", récemment introduit par Pearl, dans le cadre possibiliste. Nous montrons par la suite que les interventions peuvent être représentées de manières différentes mais équivalentes dans les réseaux causaux possibilistes. La deuxième contribution est un nouvel algorithme de propagation permettant de traiter aussi bien les observations que les interventions. Nous montrons que le coût supplémentaire pour le traitement des interventions est négligeable et que notre algorithme est plus approprié lorsqu'il s'agit de traiter des séquences d'observations et d'interventions.

Mots-clés : Réseaux possibilistes, causalité, propagation, inférence possibiliste, interventions.

1 Introduction

Les réseaux probabilistes bayésiens Pearl (1988); Jensen (1996); Lauritzen & Spiegelhalter (1988) représentent des outils de calcul puissants permettant d'identifier les interactions entre les variables à partir de données observées. Récemment, Pearl (2000) a proposé des approches basées sur la théorie de probabilité en utilisant des graphes causaux afin de formaliser la notion d'interventions. Du point de vue de la représentation, les interventions et les observations sont distinguées en utilisant le concept de l'opérateur "do" introduit par Goldszmidt & Pearl (1992); Pearl (2000). Du point de vue du raisonnement, le traitement des interventions consiste à altérer le graphe en excluant toutes les causes directes associées à la variable d'intérêt autres que les interventions en maintenant intact le reste du graphe. Les effets de telles interventions sur le reste des variables sont calculés en appliquant le conditionnement sur le graphe altéré.

L'opérateur "do" a été initialement proposé dans Goldszmidt & Pearl (1992) dans le cadre des fonctions ordinales conditionnelles de Spohn (1988b,a). Les révisions et les mises à jour sont unifiées à travers le conditionnement sur les actions afin de permettre

le raisonnement causal. Les fonctions ordinales conditionnelles de Spohn représentent un modèle "qualitatif" pour la représentation de l'incertitude et possèdent des liens très étroits aussi bien avec les probabilités infinitésimales qu'avec la théorie des possibilités (voir Dubois & Prade (1991)).

Cet article se focalise sur le développement des réseaux causaux possibilistes dans le but de traiter aussi bien les interventions que les observations. Une intervention est le résultat d'une action externe (non prévue dans le système) qui force certaines variables à prendre des valeurs données. Une observation est une nouvelle information sur la valeur de certaines variables qui complète nos croyances sur le monde réel. Une observation concerne un monde statique alors qu'une intervention concerne un monde dynamique. Cette différence ressemble à la différence entre révision et mise à jour qui est très peu étudiée dans le cadre des représentations graphiques. Deux contributions principales y sont présentées :

- Les fondements théoriques des réseaux causaux possibilistes. Plus précisément, cet article permet de définir les réseaux causaux possibilistes et introduit l'opérateur "do" à la théorie des possibilités. Nous montrons que les différentes, mais équivalentes, représentations des interventions dans les réseaux probabilistes restent valables pour le cadre de la théorie des possibilités.

- Un nouvel algorithme de propagation permettant de traiter les interventions et les observations. Nous commençons par montrer que l'adaptation directe des algorithmes de propagation dans les réseaux probabilistes n'est pas appropriée pour le traitement incrémental de séquences de nouvelles interventions et observations. En effet, pour les réseaux à connexions multiples, lorsqu'il s'agit d'interventions, répondre aux requêtes n'est plus immédiat. Ceci requiert $O(|D|^{|u| \times |r|})$ calculs où $|D|$ dénote la taille du domaine de la variable concernée par des interventions, $|u|$ est le nombre des instances de parents et $|r|$ est le nombre des différentes interventions. Si on impose l'efficacité dans le traitement des requêtes alors une solution possible consiste à utiliser les graphes augmentés (les interventions sont dans ce cas considérées comme étant des observations au niveau du graphe augmenté). Cette solution n'est pas satisfaisante dans le cadre des réseaux probabilistes du moment que, pour les graphes à connexions multiples par exemple, le traitement incrémental de séquences d'observations et d'interventions est impossible (à moins de répéter à chaque intervention une étape d'initialisation du graphe).

Cet article tire profit des différents avantages qu'offrent les propriétés des réseaux possibilistes et propose un nouvel algorithme où d'une part, l'arbre de jonction associé au réseau causal possibiliste est construit une seule fois, d'autre part, la réponse aux requêtes est réalisée en un temps linéaire.

Le reste de cet article est organisé comme suit : la section suivante présente un bref rappel de la théorie des possibilités et des réseaux possibilistes. La contrepartie possibiliste de l'opérateur do est proposée dans la troisième section. Finalement, nous présentons notre nouvel algorithme permettant de traiter les séquences non-simultanées d'interventions. La dernière section conclut cet article.

2 La théorie des possibilités

Cette section présente un bref rappel de la théorie des possibilités ; pour plus de détails voir Dubois & Prade (1988).

Soit $V = \{A_1, A_2, \dots, A_n\}$ un ensemble de variables. D_{A_i} dénote le domaine fini associé à la variable A_i . a_i dénote une instance quelconque de la variable A_i . X, Y, Z, \dots représentent des ensembles de variables. x est un élément du produit cartésien $\times_{A_i \in X} D_{A_i}$ qui est un sous-ensemble de $\Omega = \times_{A_i \in V} D_{A_i}$ l'univers de discours. ω , un élément de Ω , est appelée *interprétation* ou *événement*. Il est noté par un tuple (a_1, \dots, a_n) , où les a_i sont respectivement des instances des A_i . $\omega[A_i]$ correspond à la valeur associée à la variable A_i dans l'interprétation ω .

2.1 Mesures de possibilité et distributions de possibilité

Une distribution de possibilité π est une application de Ω vers l'intervalle $[0, 1]$. Le degré de possibilité $\pi(\omega)$ représente la compatibilité de ω avec l'information disponible. Par convention, $\pi(\omega) = 1$ signifie que ω est totalement possible, et $\pi(\omega) = 0$ signifie que ω est impossible. Lorsque $\pi(\omega) > \pi(\omega')$, ω est préférée à ω' pour être l'état réel du monde. La distribution de possibilité π est dite normalisée s'il existe au moins une interprétation cohérente avec les informations disponibles, c'est-à-dire : $\exists \omega \in \Omega, \pi(\omega) = 1$. Une mesure de possibilité Π est une fonction qui associe à chaque $\varphi \subseteq \Omega$ un poids dans l'intervalle $[0, 1]$. Π peut être simplement obtenue à partir de π comme suit : $\Pi(\varphi) = \max\{\pi(\omega) : \omega \in \varphi\}$.

Le conditionnement Dubois & Prade (1988) consiste à augmenter proportionnellement les éléments cohérents avec x :

$$\pi(\omega \mid x) = \begin{cases} \frac{\pi(\omega)}{\Pi(x)} & \text{si } \omega[X] = x \\ 0 & \text{sinon.} \end{cases} \quad (1)$$

2.2 Les réseaux possibilistes

Les réseaux possibilistes Fonck (1994); Borgelt *et al.* (1998); Ben Amor *et al.* (2003), notés par G , sont des graphes acycliques orientés (DAG). Les nœuds correspondent aux variables et les arcs représentent les relations entre les variables. Un nœud A_j est dit un parent de A_i s'il existe un lien du nœud A_j vers le nœud A_i . Les parents de A_i sont notés par U_{A_i} ou bien U_i .

Il existe deux types de réseaux possibilistes (selon le conditionnement utilisé). Cet article se focalise sur les réseaux possibilistes où le conditionnement est donné par l'équation 1.

L'incertitude au niveau des nœuds est représentée par des distributions de possibilité locale. Plus précisément, pour chaque variable A_i et pour chaque u_i un élément du produit cartésien des domaines des variables qui sont les parents de A_i , nous associons un degré de possibilité $\Pi(a_i \mid u_i)$ pour tout $a_i \in D_{A_i}$, tel que $\max_{a_i \in D_{A_i}} \pi(a_i \mid u_i) = 1$. Les réseaux possibilistes constituent des représentations compactes des distributions de possibilité. Plus précisément, les distributions de possibilité jointes associées aux réseaux possibilistes sont calculées en utilisant la règle de chaînage basée sur le produit

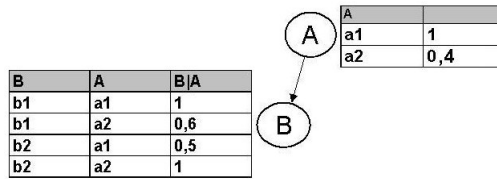


FIG. 1 – Exemple de réseau causal possibiliste G

qui est similaire à la règle de chaînage probabiliste, à savoir :

$$\pi_{\Pi G}(a_1, \dots, a_n) = \prod_{i=1, \dots, n} \Pi(a_i | u_i) \tag{2}$$

Exemple 1

La figure 1 donne un exemple des réseaux possibilistes. La distribution de possibilité jointe $\pi_G(AB) = \pi_G(A). \pi_G(B|A)$ associée à G est donnée par la table 1.

A	B	$\pi_G(AB)$
a1	b1	1
a1	b2	0.5
a2	b1	0.24
a2	b2	0.4

TAB. 1 – La distribution de possibilité jointe $\pi_G(AB)$

3 Les réseaux causaux possibilistes

La capacité de prédire l'effet des interventions dans les réseaux causaux requiert des conditions plus fortes lors de la construction de ces réseaux. Un réseau causal possibiliste est un réseau possibiliste où les arcs orientés dans le graphe représentent des relations causales entre les événements. Les arcs suivent le sens du processus causal. Intuitivement, l'ensemble de parents U_i de A_i représente la totalité des causes directes de A_i . Les arcs indiquent seulement que la variable est causalement pertinente par rapport aux autres, mais n'indique rien concernant la manière avec laquelle elle les influencerait.

La relation entre l'information causale et l'information probabiliste a fait l'objet de plusieurs recherches (notamment Pearl (2000)). Notre travail est parti de notre conviction que les mêmes résultats tiennent aussi pour les autres théories d'incertitude du moment que la transformation d'un réseau probabiliste lui permettant le traitement de la causalité concerne essentiellement la structure graphique. Dans cet article, nous nous focalisons sur les interprétations possibilistes des relations causales.

Les interventions sont traitées comme des modalités sur les variables. Une simple intervention, dite "atomique", est une intervention qui force une variable unique A_i à prendre une valeur a'_i . Cette intervention sur A_i est notée $do(A_i = a'_i)$ ou $do(a'_i)$. $do(a'_i)$ est considérée comme étant une intervention externe qui n'altère qu'un seul mécanisme

(parent-enfant) en laissant intact les autres mécanismes.

La section suivante propose un modèle possibiliste permettant de représenter les interventions en utilisant les réseaux causaux possibilistes.

3.1 L'intervention comme une négation de toutes les autres causes directes

Pearl et Verma (Verma & Pearl (1990)) interprètent les relations causales parent-enfant dans un DAG comme une fonction déterministe $a'_i = f_i(u_i, \gamma_i)$, $i = 1, \dots, n$ où u_i sont des parents de la variables A_i dans le graphe causal. γ_i , $1 \leq i \leq n$ sont des perturbations indépendantes et sont des instances de l'ensemble de variables non mesurées Γ .

L'effet d'une intervention, notée $do(A_i = a'_i)$ ou $do(a'_i)$, sur Y (un sous-ensemble de V) est déduit du modèle en supprimant les fonctions définissant A_i et en remplaçant $A_i = a'_i$ dans le reste des fonctions.

Dans les modèles graphiques, les interventions sur une variable A_i expriment aussi que nos croyances (exprimées dans un cadre incertain, ici la théorie des possibilités) sur les parents U_i de A_i doivent être affectées. Ceci est représenté par la suppression (aussi appelé mutilation) des liens entre U_i et A_i . Le reste du graphe reste intact. Le graphe mutilé résultant est noté G_m tel que $\pi(\omega|do(a'_i)) = \pi_{G_m}(\omega|a'_i)$, où π_{G_m} correspond à la distribution de possibilité associée au graphe mutilé G_m .

Le calcul de l'effet de l'intervention $do(a'_i)$ consiste à transformer la distribution $\pi(\omega)$ en une autre distribution de possibilité $\pi_{a'_i}(\omega) = \pi(\omega|do(a'_i))$. Formellement,

$$\forall \omega, \pi_{a'_i}(\omega) = \pi(\omega|do(a'_i)) = \pi_{G_m}(\omega|a'_i) \quad (3)$$

En mutilant le graphe, toutes les autres causes directes (les parents) autres que l'action deviennent indépendantes de la variable d'intérêt.

De plus, l'évènement qui associe à la variable A_i la valeur a'_i après la réalisation de l'intervention $do(a'_i)$ devient certain. Formellement, $\pi_{G_m}(a'_i) = 1$ et $\forall a_i \in D_{A_i} : a_i \neq a'_i, \pi_{G_m}(a_i) = 0$. L'effet d'une telle intervention est donc donné comme suit, $\forall \omega$:

$$\pi(\omega|do(a'_i)) = \begin{cases} \prod_{j \neq i} \pi(a_j|U_j(\omega)) & \text{si } \omega[A_i] = a'_i \\ 0 & \text{sinon} \end{cases} \quad (4)$$

où $U_j(\omega) = u_j$ correspond à la valeur que ω associe aux parents de a_j .

Exemple 2

En considérant le réseau causal possibiliste G présenté dans l'exemple 1, le graphe mutilé G_m obtenu après l'intervention $do(B = b_1) = do(b_1)$ est donné par la figure 2.

L'effet de cette intervention $do(B = b_1)$ sur la distribution jointe $\pi(AB)$ associée au graphe dans l'exemple 1 est présenté dans la table 2.

La forme de l'équation (4) est intéressante puisqu'elle permet de calculer l'effet d'une intervention $do(a'_i)$ sur un ensemble de variable Y disjoint de $\{A_i\} \cup U_i$:

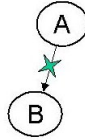


FIG. 2 – Le graphe mutilé G_m

A	B	$\pi(AB do(b_1))$
a_1	b_1	1
a_1	b_2	0
a_2	b_1	0.4
a_2	b_2	0

TAB. 2 – La distribution de possibilité jointe $\pi(AB|do(b_1))$

Proposition 1

Soit Y un ensemble de variables disjoint de $\{A_i \cup U_i\}$ où A_i est une variable dans V manipulée par une intervention $do(a'_i)$ et U_i est l'ensemble des causes directes A_i :

$$\pi(y|do(a'_i)) = \max_{u_i} \pi(y|a'_i, u_i) \cdot \pi(u_i) \tag{5}$$

En effet, il s'agit de la marginalisation sur les parents de A_i de l'expression équivalente à 4 : $\pi(a_1, \dots, a_n|do(a'_i)) = \pi(a_1, \dots, a_n|a'_i, u_i) \cdot \pi(u_i)$. La proposition 1 est la contrepartie possibiliste de la proposition de (Pearl (2000)). Ce résultat n'est pas surprenant étant donnée la similarité entre les réseaux causaux possibilistes basés sur le produit et le cas particulier des réseaux probabilistes où les probabilités conditionnelles sont soit proches de 1 soit proches de 0.

3.2 Ajout de nœuds supplémentaires

Une approche alternative mais équivalente Pearl (1993) consiste à considérer l'intervention comme étant une nouvelle variable dans le système. Cette sous section montre que cette approche alternative est aussi valable dans la théorie des possibilités. L'effet d'une intervention est calculé en appliquant le conditionnement sur le graphe altéré (augmenté). Cette altération consiste à ajouter un nœud parent DO_i à chaque variable A_i susceptible d'être concernée par une intervention. Le graphe résultant est dit augmenté. La variable DO_i prend ses valeurs dans $\{\{do_{a'_i} : \forall a'_i \in D_{A_i}\}, do_{i-noact}\}$. La valeur " $do_{i-noact}$ " (ou " $do_{A_i-noact}$ ") signifie qu'aucune intervention n'est appliquée sur A_i . Les valeurs $do_{a'_i}$ signifient que la variable A_i est forcée à prendre la valeur a'_i . Le graphe augmenté est noté G_a . L'ensemble des parents U_i est augmenté de la variable DO_i ($U'_i = U_i \cup DO_i$).

La nouvelle distribution de possibilité locale au niveau de la variable A_i après augmentation du graphe est donnée par :

$$\pi(a_i|u'_i) = \begin{cases} \pi(a_i|u_i) & \text{si } DO_i = do_{i-noact} \\ 1 & \text{si } a_i = a'_i \\ 0 & \text{si } a_i \neq a'_i \end{cases} \tag{6}$$

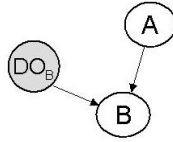


FIG. 3 – Le graphe augmenté G_a

Comme c'est le cas dans les réseaux probabilistes, les deux approches d'interprétation des interventions (mutilation et augmentation du graphe) sont aussi équivalentes dans le cadre de la théorie des possibilités. Plus précisément,

Proposition 2

Soient G un réseau causal possibiliste et G_m (resp. G_a) le graphe mutilé (resp. augmenté) obtenu à partir de G après l'application de l'intervention $do(a'_i)$ sur la variable A_i dans G . Nous avons, $\forall \omega, \forall a'_i \in D_{A_i}$,

$$\pi_{G_a}(\omega | DO_i = do_{a'_i}) = \pi_{G_m}(\omega | A_i = a'_i) = \pi(\omega | do(a'_i))$$

Cette approche a l'avantage de s'appliquer à tout type d'intervention. En ajoutant un lien $DO_i \rightarrow A_i$ à chaque nœud (sur lequel une intervention est possible) dans le graphe, il devient possible de construire une distribution de possibilité contenant des informations sur plusieurs types d'interventions.

Exemple 3

Le graphe augmenté G_a obtenu après l'intervention $do(b_1)$ à partir du graphe G introduit dans l'exemple 1 est donné par la figure 3.

La distribution de possibilité locale $\pi(B|A, DO_B)$ au niveau de B est donnée en utilisant l'équation (6). Par exemple, $\pi(b_1|a_2, do_{B-noact}) = \pi(b_1|a_2) = 0.6$ et $\pi(b_1|a_2, do_{b_2}) = 0$.

4 Propagation dans les réseaux causaux possibilistes

Dans la section précédente, nous avons montré que l'interprétation probabiliste des interventions peut facilement être adaptée au cadre de la théorie des possibilités. Cette section introduit un nouvel algorithme permettant le raisonnement à partir de réseaux causaux possibilistes. Notre algorithme est plus efficace qu'une simple adaptation de l'algorithme probabiliste. Par ailleurs, il permet le traitement de séquences non simultanées d'interventions et d'observations.

Le calcul de l'effet de séquences d'interventions et d'observations est réalisé :

- soit en généralisant la formule explicite (5) pour le traitement des ensembles d'observations $E \subseteq V$ (voir l'équation 1) et d'interventions $F \subseteq V$.
- soit en ajoutant un nœud parent à chaque variable dans F (augmenter le graphe) et appliquer par la suite le conditionnement possibiliste standard où les interventions sont traitées comme étant des observations au niveau des nœuds ajoutés.

La généralisation de (5) est donnée comme corollaire de la proposition 1.

Corollaire 1

Soit $E = e$ une observation. Soit F un ensemble de variables affectées par des interventions et f un élément de $\times_{A_i \in F} D_{A_i}$. Soit $U_F = \bigcup_{A_i \in F} U_i$, où U_i dénote l'ensemble de parents de A_i . Nous notons par u_F un élément de $\times_{A_i \in U_F} D_{A_i}$. L'effet de l'ensemble des interventions notées $do(f)$ et des observations e sur le reste des variables $A_j = a_j$ avec $A_j \notin E \cup F \cup U_F$ est donné par :

$$\Pi(a_j|e, do(f)) = \max_{u_F} \pi(a_j|e, f, u_F) \cdot \pi(u_F|e) \quad (7)$$

Ainsi, afin de calculer $\Pi(a_j|e, do(f))$ il suffit de calculer pour chaque u_F l'expression $\pi(a_j|e, f, u_F) \cdot \pi(u_F|e)$ et considérer par la suite le maximum des résultats obtenus. L'expression $\pi(a_j|e, f, u_F) \cdot \pi(u_F|e)$ est obtenue en utilisant une adaptation directe des algorithmes de propagation probabilistes puisqu'il s'agit simplement du conditionnement. Manifestement, cette solution n'est pas satisfaisante à moins que le nombre d'interventions est faible. En effet, utiliser telle équation semble inadéquat notamment pour le traitement de variables avec un nombre important de parents. Cette opération requiert $O(|D|^{|u| \times |r|})$ où $|D|$ correspond à la taille du domaine de la variable concernée par des interventions, $|u|$ est le nombre des instances de parents et $|r|$ représente le nombre d'interventions.

Un autre moyen permettant de calculer les effets des interventions consiste à mutiler le graphe initial. Cette approche est inadéquate notamment pour les graphes à connexions multiples qui impliquent une transformation graphique vers des arbres de jonction (à partir desquels une inférence efficace est effectuée). En effet, soit G un réseau possibiliste à connexions multiples. Supposons qu'une intervention $do(a'_i)$ est appliquée sur une variable A_i . Le graphe est alors mutilé et tous les liens vers A_i sont supprimés. L'arbre de jonction est formé après cette mutilation. Supposons maintenant qu'une autre intervention $do(a'_j)$ est appliquée par la suite sur A_j tel que $j \neq i$. En mutilant le graphe (déjà mutilé une fois), l'arbre de jonction résultant n'est pas le même et doit être reconstruit. Le processus de la construction de l'arbre de jonction est coûteux. Ce même processus doit être répété pour chaque modification au niveau du réseau. Dans ce qui suit, nous présentons un nouvel algorithme de propagation utilisant les graphes augmentés permettant de traiter incrémentalement les séquences d'interventions et d'observations.

4.1 Un nouvel algorithme pour les réseaux causaux possibilistes

Notre objectif consiste à proposer un nouvel algorithme tel que :

- les réponses aux requêtes sont réalisées en un temps linéaire.
- une seule transformation du graphe augmenté initial à un arbre de jonction est requise.
- les observations et les interventions sont traitées incrémentalement (c.à.d. l'arbre de jonction est mis à jour incrémentalement sans avoir à le réinitialiser).

L'idée principale de notre algorithme est la capacité d'exprimer par défaut qu'il n'y a pas d'intervention au niveau du graphe augmenté et permettre par la suite la remise en question de cet état en mettant à jour les distributions de possibilité dans le cas où quelques interventions se produisent. C'est-à-dire, il s'agit d'associer des distributions

de possibilité aux nœuds ajoutés permettant d'exprimer qu'il n'existe pas d'intervention par défaut. Pour les réseaux probabilistes, il n'existe pas de distribution remplissant cette exigence à moins de proposer de réinitialiser l'arbre de jonction pour chaque observation ou intervention. En effet, soit BN un réseau bayésien et soit F un ensemble de variables dans G susceptible d'être affectées par des interventions. A chaque nœud A_i dans F nous associons un nœud parent DO_i . Le graphe résultant est noté BN_a . Intuitivement, on s'attend à ce que la réalisation de l'inférence sans interventions et sans observations sur le graphe augmenté BN_a produise des distributions de probabilité (P_{BN_a}) équivalente à celle associée au graphe initial (P_{BN}) sur l'ensemble des variables initiales. Or, pour chaque distribution de probabilité associée aux nœuds ajoutés DO_i , y compris la distribution de probabilité uniforme, la distribution a posteriori (résultante) est différente de la distribution initiale à l'exception de la distribution qui associe le degré 1 aux valeurs " $do_{i-noact}$ " pour chaque nœud DO_i et le degré 0 aux valeurs restantes. Plus formellement, $\forall \omega, \forall i : A_i \in F$,

$$P_{BN}(\omega) = P_{BN_a}(\omega) \text{ iff } P_{BN_a}(do_{i-noact}) = 1$$

De telles distributions (le degré 1 à $do_{i-noact}$) excluent toute intervention future. Autrement dit, associer le degré 0 aux autres valeurs que " $do_{i-noact}$ " signifie qu'avoir des interventions sur A_i devient un évènement impossible.

En théorie des possibilités, la situation est différente ce qui nous a motivé à proposer notre nouvel algorithme.

Les distributions de possibilité locales pour les variables DO_i

La distribution de possibilité locale permettant d'exprimer "pas d'intervention" par défaut sans exclure la possibilité d'interventions futures est définie dans ce qui suit.

Définition 1

Soient G un réseau causal et G_a le graphe augmenté obtenu à partir de G en ajoutant un nœud parent DO_i pour chaque variable A_i dans F (l'ensemble de variables concernées par les interventions). Soit $A_i \in F$ un nœud susceptible d'être concernée par une intervention. La distribution de possibilité a priori (au départ) au niveau d'un nœud ajouté DO_i est définie par :

- $\pi_{G_a}(DO_i = do_{i-noact}) \leftarrow 1$,
- $\forall a_i \in D_{A_i}, \pi_{G_a}(DO_i = do_{a_i}) \leftarrow \epsilon$,

où ϵ est un infinitesimal (en effet, ϵ doit être tel que $\epsilon \leq \min_{\omega \in \Omega} \pi_G(\omega)$),

où π_G (resp. π_{G_a}) est la distribution de possibilité associée au graphe G (resp. G_a).

En effet, en associant le degré de possibilité 1 à la valeur " $do_{i-noact}$ " pour chaque nœud ajouté DO_i , les évènements $\{DO_i = do_{i-noact} : A_i \in F\}$ sont acceptés par défaut comme étant des situations normales. En associant le degré ϵ à $\pi(DO_i = do_{a_i})$ pour chaque instance $a_i \in D_{A_i} : A_i \in F$, les évènements $\{DO_i = do_{a_i} : A_i \in F\}$ sont considérés les moins normaux et les moins préférés dans le graphe évitant ainsi d'influencer nos connaissances initiales sur le reste des variables (les variables initiales).

La proposition suivante montre que tant qu'il n'y a ni observation ni intervention, les distributions jointes associées aussi bien pour le graphe initial que pour le graphe augmenté sont équivalentes. Plus précisément,

Proposition 3

Soit F l'ensemble des variables manipulées. G_a représente le graphe augmenté construit à partir de G en ajoutant les nœuds parents DO_i à chaque $A_i \in F$ dont les distributions locales sont données par la définition 1. La distribution de possibilité jointe π_{G_a} associée à G_a sur l'ensemble des variables initiales $V = \{A_1, \dots, A_n\}$ est équivalente à la distribution de possibilité jointe associée au graphe initial G . Formellement,

- i) $\forall \omega \forall i : 1, \dots, n, \pi_{G_a}(\omega) = \pi_G(\omega) = \pi_{G_a}(\omega | do_{i-noact})$,
- ii) $\forall \omega \forall i : 1, \dots, n, \pi_{G_a}(\omega | DO_i = do_{a_i}) = \pi_G(\omega | do(a_i))$.

Preuve 1

Soient ω une interprétation sur l'ensemble de variables $V = \{A_1, \dots, A_i, \dots, A_n\}$ et do_i une instance quelconque de la variable DO_i . Nous avons,

$$\begin{aligned}
 i) \pi_{G_a}(\omega) &= \max_{do_i} \pi_{G_a}(\omega, do_i) \\
 &= \max_{do_i} (\pi_{G_a}(a_1 | u_1) \dots \pi_{G_a}(a_i | u_i, do_i) \dots \pi_{G_a}(a_n | u_n) \cdot \pi_{G_a}(do_i)) \\
 &= [\prod_{a_j \neq a_i} \pi_{G_a}(a_j | u_j)] \cdot [\max(\pi_{G_a}(a_i | u_i, do_{i-noact}), \pi_{G_a}(do_{i-noact}), \\
 &\quad \max_{do_{a'_i}} (\pi_{G_a}(a_i | u_i, do_{a'_i}) \cdot \pi_{G_a}(do_{a'_i}))) \\
 &= [\prod_{a_j \neq a_i} \pi(a_j | u_j)] \cdot [\max(\pi_{G_a}(a_i | u_i, do_{i-noact}), \pi_{G_a}(do_{a_i}))] \\
 &= [\prod_{a_j \neq a_i} \pi(a_j | u_j)] \cdot [\max(\pi_{G_a}(a_i | u_i), \epsilon)] \\
 &= \prod_{a_i \in D_{A_i \in V}} \pi(a_i | u_i) = \pi_G(\omega)
 \end{aligned}$$

Nous avons aussi,

$$\begin{aligned}
 \pi_{G_a}(\omega | DO_i = do_{i-noact}) &= \pi_{G_a}(a_1 | u_1) \dots \pi_{G_a}(a_i | u_i, do_{i-noact}) \dots \pi_{G_a}(a_n | u_n) \\
 &= \pi_{G_a}(a_1 | u_1) \dots \pi_{G_a}(a_i | u_i) \dots \pi_{G_a}(a_n | u_n) \\
 &= \pi_{G_a}(\omega) = \pi_G(\omega)
 \end{aligned}$$

ii) Lorsque $DO_i \neq do_{i-noact}$, nous obtenons $\forall a_i : \omega[A_i] = a_i$:

$$\begin{aligned}
 \pi_{G_a}(\omega | DO_i = do_{a_i}) &= \pi_{G_a}(a_1 | u_1, do_{a_i}) \dots \pi_{G_a}(a_i | u_i, do_{a_i}) \dots \pi_{G_a}(a_n | u_n, do_{a_i}) \\
 &= \pi_{G_a}(a_1 | u_1) \dots \pi_{G_a}(a_i | u_i, do_{a_i}) \dots \pi_{G_a}(a_n | u_n)
 \end{aligned}$$

En utilisant l'équation 6 (définition de $\pi_{G_a}(a_i | u_i, do_{a_i})$), nous obtenons le même résultat que $\pi_G(\omega | do(a_i))$ (voir 4).

Lorsque $\omega[A_i] \neq a_i$,

$$\pi_{G_a}(\omega | DO_i = do_{a_i}) = \pi_G(\omega | do(a_i)) = 0.$$

Ce résultat peut facilement être étendu pour le traitement de plusieurs interventions.

Exemple 4

Considérons le réseau causal possibiliste G dans la figure 1 et le graphe augmenté G_a dans la figure 3. La distribution de possibilité locale au niveau du nœud DO_B est donnée dans la table 3. La distribution de possibilité locale au niveau de B est calculée en utilisant (6). La distribution de possibilité π_{G_a} associée à G_a est donnée dans la table 4. La distribution de possibilité π_G (voir la table 1) et π_{G_a} sur les nœuds initiaux A et B sont équivalentes. Par exemple, $\pi_G(a_2, b_1) = \pi_{G_a}(a_2, b_1) = 0.24$.

DO_B	$\pi_{G_a}(DO_B)$
$do_{B-noact}$	1
do_{b_1}	0.001
do_{b_2}	0.001

TAB. 3 – La distribution de possibilité locale $\pi_{G_a}(DO_B)$

A	B	DO_B	π_{G_a}	A	B	DO_B	π_{G_a}
a_1	b_1	$noact$	1	a_2	b_1	$noact$	0.24
a_1	b_1	do_{b_1}	0.001	a_2	b_1	do_{b_1}	0.001
a_1	b_1	do_{b_2}	0	a_2	b_1	do_{b_2}	0
a_1	b_2	$noact$	0.5	a_2	b_2	$noact$	0.4
a_1	b_2	do_{b_2}	0.001	a_2	b_2	do_{b_2}	0.001

où $noact = do_{B-noact}$

TAB. 4 – La distribution de possibilité jointe $\pi_{G_a}(A, B, DO_B)$

4.2 Les arbres de jonction augmentés

Un deuxième point important est que les nœuds " DO_i " peuvent d'une manière équivalente être ajoutés soit avant soit après la construction de l'arbre de jonction. De point de vue calculatoire, il est préférable de commencer par la construction de l'arbre de jonction à partir du réseau possibiliste initial (avant de l'augmenter). Une fois l'arbre de jonction construit, nous procédons à l'addition de nouveaux nœuds (DO_i) ainsi que des distributions qui leur sont associées ($\pi(A_i|U_i, DO_i)$ et $\pi(DO_i)$) aux distributions locales au niveau des cliques lors de l'initialisation de l'arbre de jonction.

Notre nouvel algorithme se résume aux étapes suivantes :

- Construire un arbre de jonction (JT) à partir du graphe initial G .
- Pour chaque clique $C_i, \forall \omega, \pi_{C_i}(\omega) = 1$.
- Pour chaque nœud A_i dans G , choisir une clique C_i contenant $A_i \cup U_i$
 - si $A_i \in F$, alors
 - ajouter le nœud DO_i à C_i
 - $\pi_{C_i} \leftarrow (\pi_{C_i} \cdot \pi(A_i|U_i, DO_i) \cdot \pi(DO_i))$ (utilisant l'éq. 6 et la définition 1),
 - sinon $\pi_{C_i} \leftarrow (\pi_{C_i} \cdot \pi(A_i|U_i))$.
- Intégrer l'évidence (observations et interventions).
- Propager l'évidence jusqu'à atteindre la stabilité de l'arbre de jonction.
- Répondre aux requêtes.

La proposition suivante montre que les distributions de possibilité jointes associées à l'arbre de jonction et au graphe augmenté G_a de G sont équivalentes.

Proposition 4

Soient G un réseau causal possibiliste et F un ensemble de variables manipulées dans G . do_F correspond à un élément de $\times_{A_i \in F} D_{DO_i}$ et ω est une interprétation sur $V = \{A_1, \dots, A_n\}$. Soit G_a le graphe augmenté obtenu à partir de G en ajoutant à chaque nœud A_i dans F un nœud parent DO_i . JT dénote l'arbre de jonction formé à partir de G et initialisé comme nous l'avons indiqué dans ce qui a précédé. Nous avons donc,

$$\forall \omega \forall do_F, \pi_{JT}(\omega, do_F) = \pi_{G_a}(\omega, do_F)$$

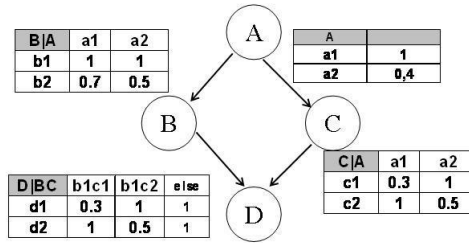


FIG. 4 – Le graphe à connexions multiples G

où π_{JT} (resp. π_{G_a}) dénote la distribution de possibilité associée à JT (resp. G_a).

Exemple 5

Considérons le réseau causal possibiliste dans la figure 4. Soient A et B des variables qui pourraient être manipulées. Le graphe augmenté résultant G_a est obtenu en ajoutant un nœud parent DO_A (resp. DO_B) à A (resp. B). Les distributions de possibilité locales associées aux nœuds DO_A et DO_B sont données comme suit : $\pi(do_{A-noact}) = \pi(do_{B-noact}) = 1$ et $\pi(do_{a_1}) = \pi(do_{a_2}) = \pi(do_{b_1}) = \pi(do_{b_2}) = 0.001$. Le degré de possibilité associée à l'évènement $\omega^+ = (a_1, b_1, c_1, d_1, do_{a_1}, do_{B-noact})$ est calculé à partir de G_a comme suit :

$$\pi_{G_a}(\omega^+) = (\pi_{G_a}(a_1|do_{a_1}).\pi_{G_a}(b_1|a_1, do_{B-noact}).\pi_{G_a}(c_1|a_1). \pi_{G_a}(d_1|b_1, c_1).\pi_{G_a}(do_{a_1}).\pi_{G_a}(do_{B-noact})) = 0.00009$$

où les distributions locales au niveau de A et B sont calculées en utilisant l'équation 6. En

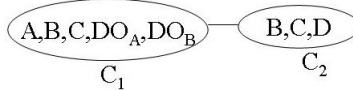


FIG. 5 – L'arbre de jonction JT après l'étape d'initialisation

initialisant l'arbre de jonction JT (formé à partir de G) consiste à initialiser les distributions de possibilité locales au niveau de chaque clique :

$$\pi_{C_1}(A, B, C, DO_A, DO_B) = (\pi(A|DO_A).\pi(DO_A). \pi(B|ADO_B).\pi(DO_B).\pi(C|A)).$$

$$\pi_{C_2}(B, C, D) = \pi(D|BC).$$

L'arbre de jonction obtenu après l'initialisation est donné dans la figure 5.

Calculer le degré de possibilité de l'évènement ω^+ à partir de l'arbre de jonction initialisé JT , nous obtenons :

$$\pi_{JT}(\omega^+) = \pi_{C_1}(\omega^+).\pi_{C_2}(\omega^+) = \pi_{C_1}(a_1, b_1, c_1, do_{a_1}, do_{B-noact}).\pi_{C_2}(b_1, c_1, d_1) = 0.00009$$

qui est égal au degré de possibilité calculé à partir du graphe augmenté G_a .

5 Conclusion

La première contribution de cet article concerne les fondements des réseaux causaux possibilistes. Nous avons montré comment les interventions peuvent être traitées en utilisant les graphes mutilés, les réseaux causaux possibilistes augmentés ainsi que les arbres de jonction augmentés. Un nouvel algorithme de propagation dans les réseaux

causaux possibilistes traitant aussi bien les observations que les interventions représente la deuxième contribution principale de cet article. Notre travail futur consiste à appliquer les résultats de cet article aux réseaux causaux possibilistes basés sur le conditionnement ordinal. Les résultats actuels montrent que les définitions de la causalité s'adaptent bien au conditionnement ordinal cependant il est difficile d'avoir la contrepartie de l'algorithme proposé dans cet article pour le cadre ordinal.

6 Remerciements

Ce travail est soutenu par le projet national ANR Blanc MICRAC.

Références

- BEN AMOR N., BENFERHAT S., DUBOIS D., MELLOULI K. & PRADE H. (2003). Anytime propagation algorithm for min-based possibilistic graphs. *Soft Computing, A fusion of foundations, methodologies and applications*, **8**, 150–161.
- BORGELT C., GEBHARDT J. & KRUSE R. (1998). Possibilistic graphical models. In *Proceedings of International School for the Synthesis of Expert Knowledge (ISSEK'98)*, p. 51–68, Udine (Italy).
- DUBOIS D. & PRADE H. (1988). *Possibility theory : An approach to computerized, Processing of uncertainty*. New York : Plenum Press.
- DUBOIS D. & PRADE H. (1991). Possibilistic logic, preferential models, non-monotonicity and related issues. In *IJCAI*, p. 419–425.
- FONCK P. (1994). *Réseaux d'inférence pour le raisonnement possibiliste*. PhD thesis, Université de Liège, Faculté des Sciences.
- GOLDSZMIDT M. & PEARL J. (1992). Rank-based systems : A simple approach to belief revision, belief update, and reasoning about evidence and actions. In *Proceeding of the Third International Conference (KR'92)*, p. 661–672. Kaufmann.
- JENSEN F. V. (1996). *Introduction to Bayesian networks*. University college, London : UCL Press.
- LAURITZEN S. L. & SPIEGELHALTER D. J. (1988). Local computations with probabilities on graphical structures and their application to expert systems. *Journal of the Royal Statistical Society*, **50**, 157–224.
- PEARL J. (1988). *Probabilistic reasoning in intelligent systems : networks of plausible inference*. San Francisco (California) : Morgan Kaufmman.
- PEARL J. (1993). Comment : Graphical models, causality and intervention. *Statistical Sciences*, **8**.
- PEARL J. (2000). *Causality : models, reasoning, and inference*. New York, NY, USA : Cambridge University Press.
- SPOHN W. (1988a). A general non-probabilistic theory of inductive reasoning. In *Proceedings of the Fourth Conference on Uncertainty in Artificial Intelligence (UAI'88)*, p. 149–158.
- SPOHN W. (1988b). Ordinal conditional functions : a dynamic theory of epistemic states causation in decision. In *Belief Changes and Statistics*, p. 105–134. W. Harper and B. Skyrms.
- VERMA T. & PEARL J. (1990). Equivalence and synthesis of causal models. In *UAI*, p. 255–270.