Cadre d'argumentation bayésien

Maxime MORGE, Julien DUMONT

morge@lifl.fr
Laboratoire d'Informatique Fondamentale de Lille
CNRS UMR 8022
Batiment M3
F-59655 VILLENEUVE D'ASCQ Cedex FRANCE
dumont@emse.fr
École Nationale Supérieure des Mines
158, cours Fauriel
F-42023 Saint-Etienne

Résumé Dans cet article, nous proposons un cadre d'argumentation bayésien. Comme n'importe quel cadre d'argumentation, il est constitué d'arguments reliés les uns aux autres par une relation de contradiction qui permet de déterminer leur acceptabilité. Afin de prendre en compte d'autres informations pertinentes, comme des préférences, un contexte ou une vraisemblance, on peut adjoindre à un cadre d'argumentation une relation de priorité entre les arguments. La plupart des cadres d'argumentation s'appuient sur un langage logique sous-jacent. Nous proposons de remplacer ce langage par un réseau bayésien. Nous sommes alors en mesure de calculer la force des arguments en fonction de la vraisemblance des cas que ces arguments mettent en exergue.

1 Introduction

L'argumentation constitue une approche prometteuse pour raisonner à partir d'informations inconsistantes. Dans [3], Dung formalise le raisonnement argumentatif à l'aide d'un cadre constitué d'un ensemble d'arguments, considérés comme des entités abstraites, reliés les uns aux autres selon une relation de contradiction qui permet de déterminer leur acceptabilité. Classiquement, les extensions de ce cadre s'appuient sur un langage logique sous-jacent [1, 4, 5]. Les arguments sont alors définis comme des schémas logiques qui relient une prémisse avec une conclusion. De manière alternative, Vreeswijk dans [6] fonde son extension sur des réseaux bayésiens en proposant un cadre d'argumentation bayésien.

Dans cet article, nous formalisons un tel cadre d'argumentation bayésien en définissant les arguments non pas comme des schémas logiques mais des arbres construits à partir de relations causales dans le réseau bayésien sous-jacent. La force des arguments est alors calculée en fonction de la vraisemblance des cas que ces arguments mettent en exergue.

Nous débuterons cet article par un bref état de l'art dans la section 2. La section 3 introduit la notion de réseau bayésien à partir de laquelle on construit notre cadre d'argumentation. Ce dernier est présenté en détail dans la section 4. Nous terminons cet article par un exemple qui illustre ce cadre (cf section 5).

2 État de l'Art

Comme dit précédemment, Dung [3] propose un cadre d'argumentation constitué d'un ensemble d'arguments reliés les uns aux autres selon une relation de contradiction qui permet de déterminer leur acceptabilité. De nombreuses extensions de ce travail ont été proposées [1, 2, 4, 5]. La plupart d'entre elles s'appuient sur un langage logique, les arguments ne sont plus de simples entités abstraites mais des schémas logiques constitués d'une prémisse qui permet de déduire une conclusion [1, 4, 5]. Nous utiliserons en lieu et place d'un tel langage logique un réseau bayésien.

Puisque d'autres informations pertinentes comme des préférences, des vraisemblances, ou un contexte, peuvent être prise en considération, la seule relation de contradiction n'est pas suffisante pour déterminer l'acceptabilité d'un argument. Par exemple, dans [1, 2, 5], une ou plusieurs relations de priorité entre arguments ont été introduites. Dans [4], la priorité affectée à un argument dépend du rôle et du contexte dans lequel l'agent impliqué se situe. Qu' elle modélise une préférence, une vraisemblance, ou un contexte, la priorité d'un argument est toujours issue d'une heuristique dépendant du domaine. S'appuyant sur un réseau bayésien, on définit un argument comme un arbre entre un ensemble de nœuds, appelés prémisses, vers une racine, appelée conclusion. Ceci nous permet de définir de manière intuitive la force des arguments selon les tables de probabilités conditionnelles sous-jacentes.

Dans la section suivante, nous définissons le réseau bayésien à partir desquels le cadre d'argumentation sera défini.

3 Réseau bayésien

Un réseau bayésien est un modèle de représentation des connaissances qui, contrairement aux approches classiques (KBase, système expert, ...), est représentable à l'aide d'un graphe. Ce dernier décrit les relations causales qui existent entre variables du problème dans un domaine particulier comme le diagnostic médical, le diagnostic de pannes ou l'analyse de risques. Ces relations de cause à effet entre les variables ne sont pas déterministes, mais probabilistes. Ainsi, l'observation d'une ou de plusieurs causes modifie la probabilité d'apparition des effets. Ce modèle de raisonnement présente l'intérêt de pouvoir prendre en compte non seulement les connaissances d'experts dans la typologie du graphe mais également l'expérience contenue dans les probabilités.

Un réseau bayésien est un graphe orienté sans cycle où chaque sommet représente une variable aléatoire et les arcs sont associés à des probabilités conditionnelles.

Definition 1 (Réseau bayésien). *On appelle réseau bayésien un graphe orienté sans cycle $BN = \langle X, V \rangle$ où :*
- $X = \{X_1, \ldots, X_n\}$ *est un ensemble de sommets qui représentent des variables aléatoires dont les domaines de valeurs* $D = \{D_1, \ldots, D_n\}$ *sont finis ;*
- V *est un ensemble d'arcs qui représentent des relations de dépendance entre couples de variables. On appelle variables parentes (dénoté* $\Pi(X_i)$*) l'ensemble des variables dont* X_i *dépend directement.*

Les relations de dépendance entre variables sont probabilistes, elles sont décrites à l'aide de **Tables de Probabilités Conditionnelles** (TPC). On associe à chaque sommet une table de probabilités conditionnelles, c'est à dire les probabilités de cette variable en fonction des valeurs possibles de ses variables parentes.

Afin de construire notre cadre d'argumentation, nous distinguons parmi les sommets d'un réseau bayésien :
- un ensemble de sommets *feuilles* (noté F) muni d'une probabilité *a priori* ;
- un ensemble de sommets factuels (noté E) parmi les sommets feuilles ($E \subseteq F$) dont nous connaissons l'état, càd dont nous considérons que la valeur est connue ;

– un ensemble de sommets d'interrogations (noté Q) dont nous souhaitons connaître la valeur. En général, $|Q| = 1$;

Puisque les probabilités conditionnelles sont explicitement représentées dans les TPC, la typologie du réseau permet d'appréhender les relations de dépendances entre les variables. La probabilité conjointe de toutes les variables du réseau est définie selon le théorème de Bayes :

$$p(X_1, \ldots, X_n) = \Pi_{i=1}^{n} p(X_i | \Pi(X_i)) \tag{1}$$

C'est à partir d'un tel réseau bayésien que nous construisons notre cadre d'argumentation.

4 Cadre d'argumentation bayésien

Lorsqu'on souhaite raisonner et interagir avec le monde, nous sommes confrontés à différentes sources d'incohérence : croyances erronées, observations non fiables, échanges d'informations contradictoires, etc ...Nous devons donc disposer d'un mécanisme de raisonnement qui permet de gérer ces incohérences. L'argumentation est un bon candidat. L'argumentation constitue un modèle de raisonnement adapté à la gestion de connaissances contradictoires.

Le cadre d'argumentation bayésien que nous présentons ici s'appuie sur les cinq notions de base suivantes :

1. **Un réseau bayésien.** Les connaissances sont des jugements factuels qui peuvent être contradictoires. Nous avons envisagé dans la section précédente des variables aléatoires reliées les unes aux autres par des relations causales.
2. **Le concept d'argument.** Un argument est : soit une entité abstraite, soit un schéma qui repose sur une base de connaissances. Dans le second cas de figure, on distingue dans un argument la prémisse de la conclusion. Un argument est en faveur d'une conclusion si la conclusion est une conséquence logique de la prémisse.
3. **La relation de contradictions entre arguments.** Puisque les connaissances peuvent être contradictoires, des arguments conflictuels peuvent co-exister. L'argumentation est une modalité naturelle pour raisonner à partir d'information contradictoire.

4. **Les relations de priorité entre arguments.** Considérer deux arguments conflictuels, n'est pas suffisant pour pouvoir choisir parmi eux, il faut prendre en considération toutes les informations pertinentes dont on dispose, notamment leur vraisemblance. Ainsi les arguments ont une force.
5. **La notion d'acceptabilité des arguments.** Les liens entre arguments permettent d'aboutir à une catégorisation des arguments appelée classes d'acceptabilité.

Un argument est constitué d'une conclusion, *i.e.* l'affectation d'une variable qui est la racine d'un arbre appelé prémisse au travers duquel la conclusion est inférée :

Definition 2. *Soit BN un réseau bayésien. Un* ***argument*** *est un couple* $A = \langle P, v_c \rangle$ *où* v_c *est une affectation de la variable* C *et* P *est un arbre de racine* C*, construit sur BN. Tous les nœuds de* P *sont instanciés et ses feuilles sont des sommets feuilles du réseau.* P *est appelée prémisse de* A*, notée* $P = premise(A)$*.* v_c *est la conclusion de* A*, dénotée par* $conclusion(A)$*.*

Un argument A' est un **sous-argument** de A si la prémisse de A' est un sous-arbre de la prémisse de A.

Puisque les valeurs des variables d'interrogation sont mutuellement exclusives, l'argumentaire ainsi construit (noté $\mathcal{A}(\mathrm{BN})$) peut être conflictuel. La relation d'attaque entre arguments, que nous instancions ici, s'appuie sur les conflits possibles entre les valeurs des variables d'interrogation qui sont mutuellement exclusives.

Definition 3. *Soient* $BN = \langle X, V \rangle$ *un réseau bayésien et* $A = \langle P_1, v_{C_1} \rangle, B = \langle P_2, v_{C_2} \rangle \in \mathcal{A}(BN)$ *deux arguments.* A ***attaque*** B *ssi* $\exists X_i \in X \; \exists v_{X_i} \in P_1 \; \exists v'_{X_i} \in P_2$ *avec* $v_{X_i} \neq v'_{X_i}$*. De même, on dit qu'un ensemble* S *d'arguments attaque* B *si* B *est attaqué par l'un des arguments de* S*.*

La force d'un argument A (noté strength(A))) correspond à la probabilité du cas mis en exergue par l'argument.

Definition 4 (Force d'un argument).
Soit BN un réseau bayésien et $A = \langle P, v_c \rangle \in \mathcal{A}(BN)$ *un argument. La* ***force de*** A *(notée strength*(A)*) correspond à la probabilité du cas,*

i.e la probabilité jointe de l'ensemble des variables de P à l'exception C :

$$strength(A) = \Pi_{X_i \in P - \{C\}} p(X_i | \Pi(X_i)) \tag{2}$$

A toute valeur d'une variable d'interrogation est associée une probabilité. Ces valeurs étant contradictoires, il existe au moins un argument en leur faveur. Afin de décider quel argument est le plus vraisemblable, on considère que l'attaque d'un argument sur un autre peut être ignorée. On introduit alors la notion de *défaite*. On dit qu'un argument défait un autre argument s'ils s'attaquent mutuellement et si le second est plus fort que le premier :

Definition 5. *Soient BN un réseau bayésien et* $A = \langle P, v_c \rangle$, $B = \langle P', v'_c \rangle \in \mathcal{A}(BN)$ *deux arguments.* A ***défait*** B *(noté* $defeats(A, B)$*) ssi :*

1. *A attaque B ;*
2. $\neg(strength(A) \leq strength(B))$.

De même, on dit qu'un ensemble S d'arguments défait B si B est défait par l'un des arguments de S.

Contrairement à la relation d'attaque, qui est symétrique et absolue, la relation de défaite est asymétrique et subjective, *i.e.* dépendante des probabilités sous-jacentes. Nous nous focalisons sur la notion d'acceptabilité suivante :

Definition 6. *Soit BN un réseau bayésien. Soient* $A \in \mathcal{A}(BN)$ *un argument et* $S \subseteq \mathcal{A}(BN)$ *un ensemble d'arguments.* A ***est acceptable vis à vis de*** S *ssi* $\forall B \in \mathcal{A}(BN)\ defeats(B, A) \Rightarrow defeats(S, B)$.

Ainsi l'ensemble des arguments acceptables constitue une position consistante, appelée extension préférée, qui peut se défendre par elle-même de toute attaque et qui ne peut pas être étendue sans introduire de conflit. Puisque la relation de priorité est une relation d'ordre, cette extension préférée est unique et non-vide [2].

L'exemple suivant permet d'illustrer notre cadre.

5 Exemple

Considérons un docteur qui délibère pour conseiller à un patient de pratiquer ou non un sport. Le réseau bayésien permettant d'élaborer un diagnostic est représenté dans la figure 1. La pratique d'un sport est requise pour les patients souffrant d'obésité et déconseillée aux personnes qui présentent un risque cardiaque. Les risques cardiaques sont d'autant plus importants que le patient est gros et vieux. L'obésité est diagnostiquée en fonction du poids et de la taille du patient. Une partie des personnes obèses le sont pour des raisons génétiques. Un tel problème génétique n'est pas systématiquement dépisté. Toutefois, on sait qu'il concerne 5% de la population.

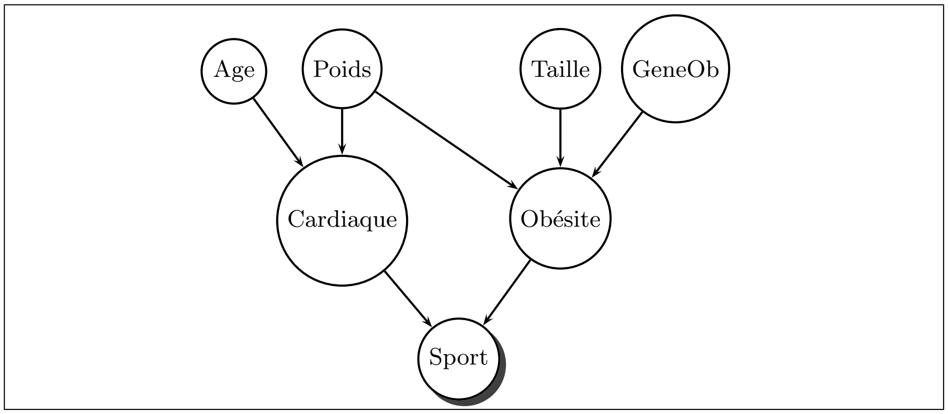

Fig. 1: Réseau bayésien de diagnostic

Les tables de probabilités conditionnelles associées aux variables Sport, Cardiac et Obesity sont respectivement représentées dans les tables 1, 2 et 3.

Le docteur considère un nouveau patient qui est jeune, grand et gros.

Les deux arguments principaux qui sont construits, l'un en faveur (P) l'autre en défaveur (C) de la pratique du sport, sont les suivants :

– $P = (\{\text{Age} = \text{jeune}, \text{Poids} = \text{gros}, \text{Taille} = \text{grand}, \text{GeneOb} = \neg\text{gene},$
$\text{Age} \rightarrow \text{Cardiaque}, \text{Poids} \rightarrow \text{Cardiaque},$
$\text{Poids} \rightarrow \text{Obésite}, \text{Taille} \rightarrow \text{Obésite}, \mathit{Obgene} \rightarrow \text{Obésite},$

```
cardiaque                    ¬cardiaque
        obese                       obese
            Pratique.1                      Pratique.8
            ¬Pratique.9                     ¬Pratique.2
        ¬obese                      ¬obese
            Pratique.1                      Pratique.6
            ¬Pratique.9                     ¬Pratique.4
```

Tab. 1. Table de probabilités conditionnelles pour la pratique du sport

```
vieux                  jeune
     gros                       gros
         cardiaque.9                cardiaque.3
         ¬cardiaque.1               ¬cardiaque.7
     maigre                     maigre
         cardiaque.3                cardiaque.1
         ¬cardiaque.7               ¬cardiaque.9
```

Tab. 2. Table de probabilités conditionnelles des risques cardiaques

```
gene                                   ¬gene
    grand           petit                  grand           petit
        gros            gros                   gros            gros
            obese.25        obese.9                obese.2         obese.8
            obese.75        ¬obese.1               ¬obese.8        ¬obese.2
        maigre          maigre                 maigre          maigre
            obese.3         obese.9                obese.1         obese.7
            ¬obese.7        ¬obese.1               ¬obese.9        ¬obese.3
```

Tab. 3. Table de probabilités conditionnelles de l'obésité

$\text{Cardiaque} \rightarrow \text{Sport}, \text{Obésite} \rightarrow \text{Sport}\}$,
$\text{Sport} = \text{Pratique})$;

– $C = (\{\text{Age} = \text{jeune}, \text{Poids} = \text{gros}, \text{Taille} = \text{grand}, \text{GeneOb} = \neg\text{gene}$,
$\text{Age} \rightarrow \text{Cardiaque}, \text{Poids} \rightarrow \text{Cardiaque}$,
$\text{Poids} \rightarrow \text{Obésite}, \text{Taille} \rightarrow \text{Obésite}, Obgene \rightarrow \text{Obésite}$,
$\text{Cardiaque} \rightarrow \text{Sport}, \text{Obésite} \rightarrow \text{Sport}\}$,
$\text{Sport} = \neg\text{Pratique})$.

Quitte à intervertir les valeurs de la variable sport, nous nous intéresserons uniquement à l'argument P. Ce dernier est constitué des deux sous-arguments suivants :

– $P_1 = (\{\text{Age} = \text{jeune}, \text{Poids} = \text{gros}$,
$\text{Age} \rightarrow \text{Cardiaque}, \text{Poids} \rightarrow \text{Cardiaque}$,
$\text{Cardiaque} \rightarrow \text{Sport}\}$,
$\text{Sport} = \text{Pratique})$;
– $P_2 = (\{\text{Poids} = \text{gros}, \text{Taille} = \text{grand}, \text{GeneOb} = \neg\text{gene}$,
$\text{Poids} \rightarrow \text{Obésite}, \text{Taille} \rightarrow \text{Obésite}, Obgene \rightarrow \text{Obésite}$,
$\text{Obésite} \rightarrow \text{Sport}\}$,
$\text{Sport} = \text{Pratique})$.

De même, l'argument P_1 est constitué des deux sous-arguments suivants :

– $P_{11} = (\{\text{Age} = \text{jeune}$,
$\text{Age} \rightarrow \text{Cardiaque}$,
$\text{Cardiaque} \rightarrow \text{Sport}\}$,
$\text{Sport} = \text{Pratique})$;
– $P_{12} = (\{\text{Poids} = \text{gros}$,
$\text{Poids} \rightarrow \text{Cardiaque}$,
$\text{Cardiaque} \rightarrow \text{Sport}\}$,
$\text{Sport} = \text{Pratique})$.

Les forces de ces différents arguments sont résumées dans le tableau 4.

On remarque qu'un argument est toujours plus fort que ses sous-arguments. Pour le patient envisagé, l'argument C est plus fort que l'argument P et ces deux arguments s'attaquent mutuellement, donc C défait P mais P ne défait pas C. Ainsi C est acceptable.

L'exemple présenté ici nous a permis d'illustrer le cadre d'argumentation bayésien.

Argument	Force
P	0.5929
C	0.4070
P_1	strength(P_1) = strength(P_{11}) $\times p$(Cardiaque = $\neg$cardiaque|Poids = gros) = .224
P_2	0.
P_{11}	strength(P_{11}) = p(Sport = Pratique|Cardiaque = $\neg$cardiaque) $\times p$(Cardiaque = $\neg$cardiaque|Age = jeune) = .56
P_{12}	p(Sport = Pratique|Cardiaque = $\neg$cardiaque) $\times p$($Cardiac$ = $\neg$cardiaque|Poids = gros) = .23

Tab. 4. Force des arguments

6 Conclusions

Dans cet article, nous avons proposé un cadre d'argumentation bayésien. Comme n'importe quel cadre d'argumentation, il est constitué d'arguments reliés les uns aux autres par une relation de contradiction qui permet de déterminer leur acceptabilité. Puisque notre cadre s'appuie sur un réseau bayésien, nous sommes en mesure de calculer la force des arguments en fonction de la vraisemblance des cas que ces arguments mettent en exergue.

Remerciements

Ce travail est cofinancé par le CPER TAC de la région Nord-Pas de Calais et les fonds européens FEDER.

Références

[1] Leila Amgoud and Claudette Cayrol. A reasoning model based on the production of acceptable arguments. *Annals of Maths and AI*, 34(1-3) :197–215, 2002.

[2] T.J.M Bench-Capon. Value based argumentation frameworks. In *Proceedings of Non Monotonic Reasoning*, pages 444–453, 2002.

[3] Phan Minh Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artif. Intell.*, 77(2) :321–357, 1995.

[4] Antonis C. Kakas and Pavlos Moraïtis. Argumentative agent deliberation, roles and context. In Jurgen Dix, João Alexandre Leite, and Ken Satoh, editors, *Electronic Notes in Theoretical Computer Science*, volume 70. Elsevier, 2002.

[5] Maxime Morge and Philippe Mathieu. Social procedure for the debate. In Benedikt Löwe Johan van Benthem, Dov Gabbay, editor, *Proc. of the 7th Augustus de Morgan Workshop*, King's college, London, 2006. to appear.

[6] G.A.W. Vreeswijk. Argumentation in bayesian belief networks. In C. Reed In I. Rahwan, P. Moraitis, editor, *First International Workshop on Argumentation in Multi-Agent Systems*, 2004.