

Chapitre 4 : Raisonnement probabiliste et possibiliste



Pr. Mustapha BOURAHLA, Département
d'Informatique, Université de M'Sila, Contact :
mustapha.bourahla@univ-msila.dz

Table des matières



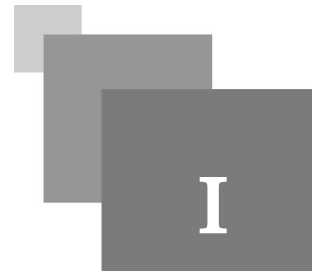
Introduction	3
I - Probabilités	4
II - Raisonnement probabiliste	7
III - Inférence par énumération	9
IV - Le langage ProbLog	12
V - Exercices sur le quatrième chapitre	15
1. Exercice : Exercice numéro 1	15
2. Exercice : Exercice numéro 2	15
3. Exercice : Exercice numéro 3	15
Conclusion	17
Solutions des exercices	18

Introduction



Ce quatrième chapitre présente le raisonnement probabiliste et possibiliste.

Probabilités



Notions de base

- Exprime le degré de croyance.
- Commencer avec un ensemble Ω - espace d'échantillonnage
 - Exemple: 6 possibilités si on roule un dé.
 - $\omega \in \Omega$ est un échantillon (un état ou un événement atomique).
- L'espace/modèle de probabilités est l'espace d'échantillonnage avec une distribution de probabilité $P(\omega)$ pour chaque élément $\omega \in \Omega$.
 - $0 \leq P(\omega) \leq 1$
 - $\sum_{\omega} P(\omega) = 1$
- Par exemple: $P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = 1/6$
- Un événement est un sous-ensemble de Ω .
- Probabilité d'un événement (c.à-d., d'un élément de Ω)
 - $P(A) = \sum_{\omega \in A} P(\omega)$
- Exemple:

Variable aléatoire

- Une variable aléatoire est une variable décrivant une partie des connaissances incertaines. Chaque variable a un domaine de valeurs qu'elle peut prendre (ex.: binaire, entier, réelle, états, etc.).
- On peut voir une variable comme une fonction définie sur l'espace d'échantillonnage et donnant une valeur à chaque échantillon en entrée:
- Types de variables aléatoires:
 1. Booléennes : le domaine est <true, false>. Exemple: .
 2. Discrètes: le domaine est énumérable. Exemple: , avec $v \in \{\text{soleil, pluie, nuageux, neige}\}$
 3. Continues: le domaine continu (par exemple, l'ensemble des réels). Exemples: $x \leq 4.02$, $position \leq 4.02$, $speed \leq 20.02$
- P induit une distribution de probabilité pour chaque variable aléatoire X :

$$P(X = x_i) = \sum_{\omega: X(\omega)=x_i} P(\omega)$$
- Exemple: Lancer un dé: $P(\text{nombre impair}) = P(1) + P(3) + P(5) = 1/6 + 1/6 + 1/6 = 1/2$
- Un événement atomique est une spécification complète de l'état pour lequel l'agent est incertain. Par exemple, si le «monde» de l'agent est décrit par seulement deux variables aléatoires booléennes (*Carie* et *MalDeDents*), il y a exactement quatre états/événements atomiques possibles:

- $Carie = false \wedge MalDeDents = false$
- $Carie = false \wedge MalDeDents = true$
- $Carie = true \wedge MalDeDents = false$
- $Carie = true \wedge MalDeDents = true$

- Les événements atomiques sont exhaustifs et mutuellement exclusifs.

Axiomes de la théorie des probabilités: Axiomes de Kolmogorov

Pour toute propositions A, B

- $0 \leq P(A) \leq 1$
- $P(true) = 1$ et $P(false) = 0$
- $P(A \vee B) = P(A) + P(B) - P(A \wedge B)$

Probabilité à priori/inconditionnelle

- La probabilité à priori ou inconditionnelle de propositions exprime le degré de croyance dans ces propositions avant l'acquisition de toute (nouvelle) information /évidence.

Exemple: $P(Carie = true) = 0.1$ et

- La distribution des probabilités donne les valeurs de probabilités pour toutes les assignations possibles de valeurs aux variables:

Exemple:

Fonction de densité de probabilités: distribution de probabilité pour une variable continue.

- La distribution conjointe de probabilités pour un ensemble de variables donne la probabilité pour chaque événement atomique décrit par ces variables.

Exemple, la distribution conjointe est une matrice 4×2 :

Météo =	soleil	pluie	nuageux	neige
Carie = true	0.144	0.02	0.016	0.02
Carie = false	0.576	0.08	0.064	0.08

Probabilité à posteriori/conditionnelle

- La probabilité conditionnelle ou à posteriori tient compte des nouvelles informations/évidences disponibles.

Exemple: $P(carie|MalDeDents) = 0.8$

C.-à-d., étant donné que la seule chose que je sais est MalDeDents.

Si on constate qu'un patient a mal aux dents et aucune autre information n'est encore disponible, la probabilité qu'il ait une carie est de 0.8.

- Si on en apprend plus, (par exemple, on découvre une carie), on a:

$P(carie | MalDeDents, carie) = 1$

- Toutes les nouvelles informations ne sont pas pertinentes, donc on peut simplifier:

Exemple:

- Ce type d'inférence, supportée par les connaissances du domaine, est crucial.

Probabilité à posteriori/conditionnelle

- Définition de la probabilité conditionnelle: $P(a|b) = P(a \wedge b)/P(b)$ si $P(b) \neq 0$ La probabilité de a , étant donné (que tout ce qu'on sait est) b .
 - Formulation équivalente (règle du produit): $P(a \wedge b) = P(a|b)P(b) = P(b|a)P(a)$
 - Il existe une version plus générale pour les distributions de probabilité. Exemple:
 - La règle suivante (chain rule) est obtenue par une application successive de la règle du produit:
- $$P(X_1, \dots, X_n) = P(X_1, \dots, X_{n-1})P(X_n|X_1, \dots, X_{n-1}) = P(X_1, \dots, X_{n-2})P(X_{n-1}|X_1, \dots, X_{n-2})P(X_n|X_1, \dots, X_{n-1}) = \dots = \pi_{i=1..n} P(X_i|X_1, \dots, X_{i-1})$$

Raisonnement probabiliste


 II

Introduction au raisonnement probabiliste

- Raisonnement avec incertitude
- Théorie des probabilités: syntaxe et sémantique
- Inférences simples
- Indépendance entre des variables aléatoires
- Règle de Bayes
- Illustration avec le monde des wumpus

Incertitude

- Soit A_t l'action d'aller à l'aéroport t minutes avant le départ de l'avion. A_t me permettra-t-il d'arriver à temps?
- Problèmes:
 - Observabilité partielle (conditions routières, etc.)
 - Senseurs bruités (annonces du trafic, etc.)
 - Incertitude dans l'effet des actions (crevaisons, pannes, etc.)
 - Immense complexité pour modéliser les actions et le trafic.
- Un raisonnement purement logique
 - Risque de tirer des conclusions erronées: « A_{25} me permettra d'arriver à temps », ou
 - Risque de tirer des conclusions peu exploitable du point de vue « prise de décision »:
 - « A_{25} me permettra d'arriver à temps, s'il ne pleut pas, s'il n'y a pas d'accident, si mes pneus ne crèvent pas, etc. »
 - « A_{1440} me permettra raisonnable d'arriver à temps, mais je devrai passer une nuit à l'aéroport. »

Probabilités pour le raisonnement avec incertitude

- Modélise la croyance/certitude des agents: Les connaissances de l'agent peuvent au mieux donner un degré de croyance dans les faits.
- Étant donné l'information/évidence disponible jusqu'ici, A_{25} me permettra d'arriver avec une probabilité de 0.4
- Probabilité subjective/bayésienne: Les probabilités expriment le degré de croyance d'un agent dans des propositions/faits. Par exemple:

- Les probabilités des propositions changent avec l'acquisition de nouvelles informations. Par exemple:



Inférence par énumération

- Commencer avec la distribution conjointe des probabilités:

	MalDeDents		¬ MalDeDents	
	Croche	¬Croche	Croche	¬Croche
Carie	0.108	0.012	0.072	0.008
¬ Carie	0.016	0.064	0.144	0.576

- Pour chaque proposition ϕ , faire une somme sur les événements atomiques pour lesquels elle est vraie:

- Commencer avec la distribution conjointe des probabilités:

	MalDeDents		¬ MalDeDents	
	Croche	¬Croche	Croche	¬Croche
Carie	0.108	0.012	0.072	0.008
¬ Carie	0.016	0.064	0.144	0.576

- Pour chaque proposition ϕ , faire une somme sur les événements atomiques pour lesquels elle est vraie :

$$P(\phi) = \sum_{\omega:\omega\models\phi} P(\omega) P(MalDeDents) = 0.108 + 0.012 + 0.016 + 0.064 = 0.2$$

- Commencer avec la distribution conjointe des probabilités:

	MalDeDents		¬ MalDeDents	
	Croche	¬Croche	Croche	¬Croche
Carie	0.108	0.012	0.072	0.008
¬ Carie	0.016	0.064	0.144	0.576

- Pour chaque proposition ϕ , faire une somme sur les événements atomiques pour lesquels elle est vraie :

$$P(\phi) = \sum_{\omega:\omega\models\phi} P(\omega) P(carie \vee MalDeDents) = 0.108 + 0.012 + 0.072 + 0.008 + 0.016 + 0.064 = 0.28$$

- Commencer avec la distribution conjointe des probabilités:

	MalDeDents		¬ MalDeDents	
	Croche	¬Croche	Croche	¬Croche
Carie	0.108	0.012	0.072	0.008
¬ Carie	0.016	0.064	0.144	0.576

- On peut aussi calculer les probabilités conditionnelles:

$$P(\neg carie | MalDeDents) = \frac{P(\neg carie \wedge MalDeDents)}{P(MalDeDents)} = (0.016 + 0.064) / (0.108 + 0.012 + 0.016 + 0.064) = 0.4$$

Normalisation

	MalDeDents		¬ MalDeDents	
	Croche	¬Croche	Croche	¬Croche
Carie	0.108	0.012	0.072	0.008
¬ Carie	0.016	0.064	0.144	0.576

Le dénominateur peut être vu comme constante de normalisation α

$$P(\text{Carie}|\text{MalDeDents}) = \alpha P(\text{Carie}, \text{MalDeDents})$$

$$= \alpha [P(\text{Carie}, \text{MalDeDents}, \text{croche}) + P(\text{Carie}, \text{MalDeDents}, \neg \text{croche})]$$

$$= \alpha \langle 0.12, 0.08 \rangle = \langle 0.6, 0.4 \rangle$$

$$\text{avec } \alpha = 1/P(\text{MalDeDents}) = 1/(\cdot 108 + \cdot 012 + \cdot 016 + \cdot 064) = 1/0.2 = 5.$$

Idée générale: Calculer la contribution à la variable d'interrogation en fixant les variables d'évidences (observées) et en faisant la somme sur les variables cachées.

En général on veut calculer la probabilité conjointe à posteriori sur un ensemble de variables d'interrogations X étant donné les valeurs e pour les variables d'évidence E .

Soit Y l'ensemble des variables cachées (non encore observées), X la valeur recherchée, et E l'ensemble des variables d'évidence.

On obtient la probabilité pour l'interrogation $P(X|E = e)$ en faisant une sommation sur les variables cachées: $P(X|E = e) = \alpha P(X, E = e) = \alpha \sum_y P(X, e, y)$

Les termes dans la somme sont des probabilités jointes vu que X, E et H pris ensembles couvrent toutes les variables aléatoires.

Défis:

- Complexité en temps: $O(d^n)$, avec d la taille du plus grand domaine des variables.
- Complexité en espace: $O(d^n)$, pour stocker la distribution.
- Comment trouver les valeurs de chacune des $O(d^n)$ entrées?

Indépendance

Les variables A et B sont indépendantes si et seulement si $P(A|B) = P(A)$ ou $P(B|A) = P(B)$ ou $P(A, B) = P(A)P(B)$ par exemple,

Indépendance conditionnelle

- Si j'ai une carie, la probabilité que la sonde croche dans la dent ne dépend pas du fait que j'aie mal à la dent ou non:
 - (1) $P(\text{Croche}|\text{MalDeDents}, \text{Carie}) = P(\text{Croche}|\text{Carie})$
- Même chose si je n'ai pas la carie:
 - (2) $P(\text{Croche}|\text{MalDeDents}, \neg \text{Carie}) = P(\text{Croche}|\neg \text{Carie})$
- croche est conditionnellement indépendant de MalDeDents étant donné carie:
 - $P(\text{Croche}|\text{MalDeDents}, \text{Carie}) = P(\text{Croche}|\text{Carie})$

- Formulations équivalentes:
 - $P(\text{MalDeDents}|\text{Croche}, \text{Carie}) = P(\text{MalDeDents}|\text{Carie})$
 - $P(\text{MalDeDents}, \text{Croche}|\text{Carie}) = P(\text{MalDeDents}|\text{Carie})P(\text{Croche}|\text{Carie})$
- Réécrivons la distribution conjointe en utilisant la règle en chaîne (chain rule):

$$P(\text{MalDeDents}, \text{Croche}, \text{Carie})$$

$$= P(\text{MalDeDents}|\text{Croche}, \text{Carie})P(\text{Croche}, \text{Carie})$$

$$= P(\text{MalDeDents}|\text{Croche}, \text{Carie})P(\text{Croche}|\text{Carie})P(\text{Carie})$$

$$= P(\text{MalDeDents}|\text{Carie})P(\text{Croche}|\text{Carie})P(\text{Carie})$$

C-à-d., $2 + 2 + 1 = 5$ nombre indépendants
- Dans des cas idéals, l'exploitation de l'indépendance conditionnelle réduit la complexité représentation de la distribution conjointe d'une exponentielle ($O(2^n)$) en linéaire ($O(n)$).
- En raisonnement probabiliste, l'indépendance conditionnelle est le concept de représentation des connaissances le plus basique et robuste.

Règle de Bayes

- Règle du produit: $P(a \wedge b) = P(a|b)P(b) = P(b|a)P(a)$
 - Règle de Bayes: $P(a | b) = P(b | a) P(a) / P(b)$
- ou, pour les distributions,
 - $P(Y|X) = P(X|Y)P(Y)/P(X) = \alpha P(X|Y)P(Y)$
- Utile pour calculer/interroger une probabilité diagnostique à partir d'une probabilité causale:
 - $P(\text{Cause}|\text{Effect}) = P(\text{Effect}|\text{Cause}) P(\text{Cause}) / P(\text{Effect})$

Exemple

Soit m (méningite), s(stiff neck / nuque raide), $P(s|m) = 0.5$, $P(m) = 1/50000$ et $P(s) = 1/20$.

- $P(m|s) = P(s|m)P(m)/P(s) = 0.5 \times 0.00002/0.05 = 0.0002$
- Règle diagnostique: Effets observés \rightarrow causes cachées
- Règle causale: causes cachées \rightarrow effets observées

Règle de Bayes et indépendance conditionnelle

$$P(\text{Carie}|\text{MalDeDents} \wedge \text{Croche}) = \alpha P(\text{MalDeDents} \wedge \text{Croche}|\text{Carie})P(\text{Carie}) = \alpha P(\text{MalDeDents}|\text{Carie})P(\text{Croche}|\text{Carie})P(\text{Carie})$$

Exemple d'un modèle Bayes simple (Bayesian classifier):

$$P(\text{Cause}, \text{Effect}_1, \dots, \text{Effect}_n) = P(\text{Cause})\pi_i P(\text{Effect}_i|\text{Cause})$$

Le langage ProbLog



IV

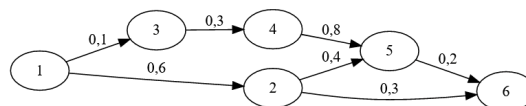
ProbLog

- ProbLog est une extension probabiliste de Prolog (la plus simple extension probabiliste de Prolog que l'on peut concevoir, ProbLog est essentiellement Prolog où toutes les clauses sont étiquetées avec la probabilité qu'elles soient vraies).
- Un programme ProbLog définit une distribution sur les programmes logiques en spécifiant pour chaque clause la probabilité d'appartenir à un programme échantillonné aléatoirement
- Ces probabilités sont mutuellement indépendantes.
- La sémantique de ProbLog est alors définie par la probabilité de succès d'une requête, qui correspond à la probabilité que la requête aboutisse dans un programme échantillonné aléatoirement.
- Un programme ProbLog spécifie une distribution de probabilité sur tous les sous-programmes non probabilistes possibles du Programme ProbLog.
- La probabilité de succès d'une requête est alors défini simplement comme la probabilité qu'il réussisse ces sous-programmes.
- ProbLog a été motivé par l'application de l'extraction de grands réseaux biologiques où les bords sont étiquetés avec des probabilités.
- Ces réseaux de biologie ont les concepts (gènes, protéines, phénotypes, etc.) qui peuvent être extraits à partir de grandes bases de données publiques et des liens probabilistes entre les concepts peuvent être obtenus par diverses techniques de prédiction.
- Dans ceci et beaucoup d'autres applications, les liens probabilistes sont mutuellement indépendants et peuvent facilement être décrits dans ProbLog.

- La probabilité de succès d'une requête ProbLog peut être calculée comme la probabilité d'une formule booléenne monotone DNF (Disjunctive Normal Form) de variables aléatoires binaires. Ce problème est NP-difficile.
- l'algorithme d'approximation de ProbLog est capable de gérer formules contenant jusqu'à 100 000 propositions.
- Le solveur ProbLog a été motivé par et utilise les progrès récents dans les diagrammes de décision binaires (BDD) pour traiter les fonctions booléennes.
- En même temps, il utilise un algorithme d'approximation pour calculer la probabilité de succès le long des lignes
- En utilisant cet algorithme, nous rapportons des expériences dans des réseaux biologiques qui démontrent l'utilité de l'approche.
- De toute évidence, il est simple de transférer ProbLog vers d'autres domaines d'exploration de liens et de réseaux.

Exemple : Graphiques probabilistes

ProbLog peut être utilisé pour modéliser des graphes probabilistes, c'est-à-dire des graphes dans lesquels l'existence de certaines arêtes est incertaine. Nous pouvons utiliser ProbLog pour calculer, par exemple, la probabilité qu'il y ait un chemin entre deux nœuds donnés.



Le programme suivant illustre cela pour un graphe à six nœuds et sept arêtes probabilistes. Le programme utilise la définition Prolog habituelle du chemin.

```

1 0.6::edge(1,2).
2 0.1::edge(1,3).
3 0.4::edge(2,5).
4 0.3::edge(2,6).
5 0.3::edge(3,4).
6 0.8::edge(4,5).
7 0.2::edge(5,6).
8
9 path(X,Y) :- edge(X,Y).
10 path(X,Y) :- edge(X,Z),
11     Y \== Z,
12     path(Z,Y).
13
14 query(path(1,5)).
15 query(path(1,6)).

```

ProbLog calcula la probabilité de l'existence d'un chemin de nœud 1 vers le nœud 5, de nœud 1 vers le nœud 6. Nous obtenons. Nous obtenons $P(\text{path}(1,5)) = 0.25824$ et $P(\text{path}(1,6)) = 0.2167296$.

Exemple

Le programme ProbLog ci-dessous encode une variante du célèbre problème Friends & Smokers. Les deux premières règles stipulent qu'il existe deux causes possibles pour une personne X de fumer, à savoir X ayant du stress, et X ayant un ami Y qui fume lui-même et influence X . Notez que, selon le programme, la présence de stress et des amis fumeurs (éventuellement multiples) contribuent tous à la probabilité que la personne X fume. De plus, le programme code que si X fume, il souffre d'asthme avec une probabilité de 0,4.

Le programme ci-dessous considère un scénario avec 4 personnes, ayant un total de 5 relations d'amitié.

```

1 0.3::stress(X) :- person(X).
2 0.2::influences(X,Y) :- person(X), person(Y).
3
4 smokes(X) :- stress(X).
5 smokes(X) :- friend(X,Y), influences(Y,X), smokes(Y).
6
7 0.4::asthma(X) :- smokes(X).
8
9 person(1).
10 person(2).
11 person(3).
12 person(4).
13
14 friend(1,2).
15 friend(2,1).

```

```

16 friend(2,4).
17 friend(3,2).
18 friend(4,2).
19
20 evidence(smokes(2),true).
21 evidence(influences(4,2),false).
22
23 query(smokes(1)).
24 query(smokes(3)).
25 query(smokes(4)).
26 query(asthma(1)).
27 query(asthma(2)).
28 query(asthma(3)).
29 query(asthma(4)).

```

1 Query	Location	Probability
2 <code>asthma(1)</code>	26:7	0.20350877
3 <code>asthma(2)</code>	27:7	0.4
4 <code>asthma(3)</code>	28:7	0.176
5 <code>asthma(4)</code>	29:7	0.176
6 <code>smokes(1)</code>	23:7	0.50877193
7 <code>smokes(3)</code>	24:7	0.44
8 <code>smokes(4)</code>	25:7	0.44

ProbLog calcule la probabilité de différentes personnes fumant et ayant asthme étant donné l'évidence que la personne 2 ne fume pas et que la personne 4 n'influence pas la personne 2.

Nous obtenons $P(\text{smoke}(1)) = 0.50877$, $P(\text{smoke}(3)) = 0.44$, $P(\text{smoke}(4)) = 0.44$,
 $P(\text{asthma}(1)) = 0.203508$, $P(\text{asthma}(2)) = 0.4$, e t
 $P(\text{asthma}(3)) = P(\text{asthma}(4)) = 0.176$.

Exercices sur le quatrième chapitre



1. Exercice : Exercice numéro 1

Soit la distribution de la probabilité suivante

	MalDeDents		¬ MalDeDents	
	Croche	¬Croche	Croche	¬Croche
Carie	0.108	0.012	0.072	0.008
¬ Carie	0.016	0.064	0.144	0.576

Question 1

[solution n°1 p.18]

Trouver la constante de normalisation α pour la probabilité $P(\text{MalDeDents})$

Question 2

[solution n°2 p.18]

Puis, calculer la probabilité $P(\text{Carie} \mid \text{MalDeDents})$

2. Exercice : Exercice numéro 2

Nous voulons calculer la probabilité de la cause d'un feu dans la forêt.

Si la probabilité d'avoir le feu est $P(\text{feu}) = 0.2$ et la probabilité d'avoir une cigarette $P(\text{cigarette}) = 0.4$

Si la probabilité de l'effet observé $P(\text{cigarette} \mid \text{feu}) = 0.6$

Question

[solution n°3 p.18]

Trouver la probabilité de la cause cachée

3. Exercice : Exercice numéro 3

Écrire un programme ProbLog pour spécifier

1. La probabilité d'aller de A à B est 0.2
2. La probabilité d'aller de B à C est 0.7.
3. La probabilité d'aller de A à C est 0.6.
4. La probabilité d'aller de A à D est 0.5
5. La probabilité d'aller de D à C est 0.2

Question 1

[solution n°4 p.18]

Le programme ProbLog

Question 2

[solution n°5 p.18]

Écrire une requête pour calculer la probabilité d'aller de A à C

Question 3

[solution n°6 p.18]

Calculer manuellement cette probabilité

Conclusion



Le raisonnement probabiliste et possibiliste ainsi qu'un langage probabiliste logique basé sur ces raisonnements sont présentés dans ce chapitre.



Solutions des exercices



> Solution n°1

Exercice p. 15

$$\alpha = 1/(0.108 + 0.012 + 0.016 + 0.064) = 5$$

> Solution n°2

Exercice p. 15

$$P(\text{Carie} \mid \text{MalDeDents}) = \alpha \times P(\text{Carie} \wedge \text{MalDeDents}) = 5 \times (0.108 + 0.012) = 0.6$$

> Solution n°3

Exercice p. 15

$$P(\text{feu} \mid \text{cigarette}) = P(\text{cigarette} \mid \text{feu}) \times P(\text{feu}) / P(\text{cigarette}) = 0.6 \times 0.2 / 0.4 = 0.12 / 0.4 = 0.3$$

> Solution n°4

Exercice p. 15

0.2 : :chemin(a,b).

0.7 : :chemin(b,c).

0.6 : :chemin(a,c).

0.5 : :chemin(a,d).

0.2 : :chemin(d,c).

chemin(X,Y) :- chemin(X,Z), chemin(Z,Y), Z \= Y.

> Solution n°5

Exercice p. 16

query(a,c)

> Solution n°6

Exercice p. 16

P(A,C) =