

Résumé : Réseaux bayésiens

D. Cram, M. May, R. Guelton, S. Touch

[damien.cram, remi.guelton]@ecl2006.ec-lyon.fr, [may.madeth,touch.sereysethy]@itc.edu.kh

Novembre, 2005

1 Introduction

L'un des enjeux principaux dans le domaine de la recherche en Intelligence Artificielle est d'être capable de concevoir et de développer des systèmes dynamiques et évolutifs. De ce fait, ces derniers doivent être équipés de comportements intelligents qui peuvent apprendre et raisonner.

Mais dans la plupart des cas, la connaissance acquise n'est pas toujours adéquate pour permettre au système de prendre la décision la plus appropriée. Pour répondre à ce genre de questions, plusieurs méthodologies ont été proposées, mais seules les approches probabilistes s'adaptent mieux non seulement au raisonnement avec la connaissance et la croyance incertaine, mais aussi à la structure de la représentation de la connaissance. Ces approches probabilistes sont appelées "réseaux bayésiens" [1][2], mais sont aussi connues sous le nom de "belief networks", "causal networks".

Les réseaux bayésiens sont la combinaison des approches probabilistes et de la théorie de graphes. Autrement dit, ce sont des modèles qui permettent de représenter des situations de raisonnement probabiliste à partir de connaissances incertaines. Ils sont une représentation efficace pour les calculs d'une distribution de probabilités [3].

2 Réseaux bayésiens

2.1 Probabilité conditionnelle

Etant donné que l'événement B s'est produit, la probabilité que l'événement A se produise est x et s'écrit : $P(A|B) = x$.

2.2 Théorème de Bayes

Les réseaux bayésiens sont initialement connus comme « réseaux de relations », mais sont nommés "Bayes" après la mort de Révérend Thomas Bayes, 1702-1761, théologien et mathématicien britannique qui a écrit une loi de base de probabilité, maintenant appelée le théorème de Bayes[4].

Etant donnés deux événement A et B qui sont conditionnellement dépendants et un contexte c , le théorème de Bayes peut s'énoncer comme ci-dessous :

$$P(B|A,c) = \frac{P(B|c)P(A|B,c)}{P(A|c)} \quad \text{où}$$

$P(B|A,c)$ est la probabilité *a posteriori* ou la probabilité de B après avoir pris en compte l'effet de A dans un contexte c ,

$P(B|c)$ est la probabilité *a priori* de l'événement B , dans le contexte c ,

$P(A|B,c)$ est la probabilité de A si l'on suppose que B est vrai dans un contexte c . Elle est appelée aussi la "vraisemblance",

$P(A|c)$ est la normalisation, c'est la probabilité de l'événement A dans le contexte c .

2.3 Réseaux bayésiens

Un réseau bayésien est un graphe acyclique orienté¹ *i.e.* c'est un graphe orienté sans circuit possible. Chaque nœud d'un réseau bayésien porte une étiquette qui est un des attributs du problème. Ces

¹ On dit souvent: un DAG. De l'anglais *directed acyclic graph*.

attributs sont binaires, pouvant prendre (avec certaine probabilité) la valeur *VRAI* ou *FAUX*, ce qui signifie qu'une variable aléatoire est associée à chaque attribut.

Sur la figure 1, on voit que *A* et *C* sont conditionnellement indépendants, on peut donc dire que $P(A|B,C) = P(A|B)$, i.e. la probabilité de *A* ne dépend que de *B*.

Autrement dit, dans le réseau bayésien tout nœud est conditionnellement indépendant de ses non-descendants sachant ses parents [3]. La probabilité de *C* est donc $P(C|B,A) = P(C|B)$.

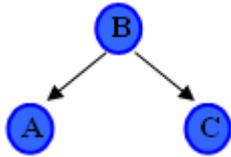


Figure 1 Un réseau bayésien

Dans l'exemple ci-contre, la probabilité jointe de toutes les variables est :

$$P(A,B,C) = P(A|B) \times P(B) \times P(C|B)$$

En général, étant donné les nœuds $X = X_1, \dots, X_n$, la probabilité jointe pour les réseaux bayésiens est :

$$P(X) = \prod_{i=1}^n P(X_i | \text{parents}(X_i)).$$

La probabilité jointe de toutes les variables est le produit de toutes les probabilités de chaque variable et ainsi les valeurs de ses parents.

2.4 La notion de causalité

Dans les réseaux bayésiens, deux nœuds qui représentent deux faits différents peuvent être en relation causale sans que l'un implique l'autre.

Graphiquement, chaque parent a un effet sur ses fils. La notion de causalité joue un rôle très important pour construire les réseaux bayésiens [3].

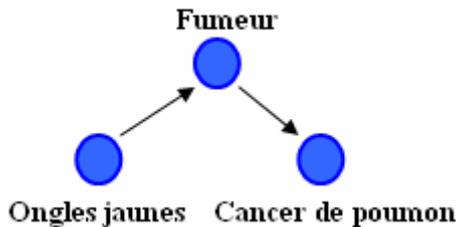


Figure 2 Graphe de causalité

Par exemple [5], sur un patient, si on constate qu'il a des ongles jaunes, alors on pourrait dire qu'il fume et en même temps, s'il fume, il a une forte chance d'avoir le cancer de poumon. Cela peut être représenté sous forme d'un graphe de causalité (voir figure 2).

2.5 Les inférences

Une fois qu'un réseau bayésien a été construit pour rendre compte d'un domaine, on cherche souvent à l'utiliser pour déterminer des probabilités correspondant à certains événements, certaines questions, certaines dépendances [3]. Dans ces cas, les inférences permettent effectivement de calculer la probabilité de chaque nœud lorsque les autres variables sont inconnues [6]. Il existe cependant quelques méthodes permettant d'effectuer ce calcul telles que *Conditionnement global*, *Arbre de jonction*, *Méthode approchée*.

2.6 Apprentissage

Lors de la construction de réseaux bayésiens, il n'est pas toujours évident qu'un expert puisse fournir de façon numérique l'ensemble des paramètres nécessaires à l'inférence dans un graphe. Il est intéressant dans certains cas de déterminer ces paramètres à partir d'une base d'exemples. C'est grâce à la méthode d'apprentissage qui permet de remédier ce problème. Les deux méthodes utilisées sont *apprentissage de paramètres* et *apprentissage de structure*.

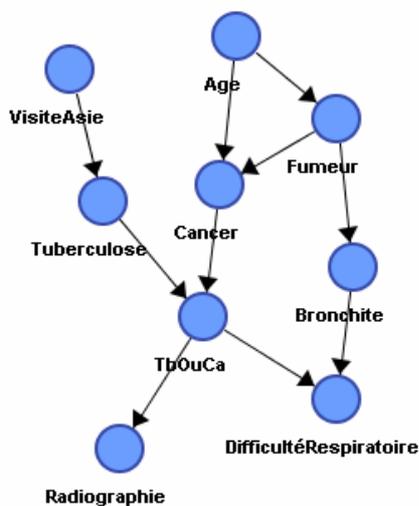
3 Les applications

3.1 Domaines d'applications

Les réseaux bayésiens sont utilisés pour le diagnostic et le raisonnement dans des systèmes complexes dans de nombreux domaines :

- **Marketing/Finance (modélisation de risques) :**
 - Détection de fraudes
 - Optimisation de portefeuilles
- **Informatique :**
 - Microsoft : différents assistants de résolution de problèmes, l'assistant Office
 - Evaluation de la qualité pour des développements logiciels
 - Reconnaissance de la parole, d'objets 3D...
 - Diagnostic de pannes
- **Médecine :**
 - Aide au diagnostic de problèmes précis,
 - Surveillance transfusionnelle
- **Industrie :**
 - NASA Vista : aide au diagnostic de pannes en temps réel pour les systèmes de propulsion
 - Lockheed Martin : système de contrôle d'un véhicule sous-marin autonome
 - Ricoh : aide au télédiagnostic
 - EDF : modélisation de groupes électrogènes
- **Environnement :**
 - Prédiction des pics d'ozone
 - Prévisions météo

3.2 Exemple d'aide à la décision : Modélisation des connaissances d'un médecin spécialiste des maladies pulmonaires



Pour mieux comprendre en quoi les réseaux bayésiens constituent un bon outil d'aide à la décision, voyons sur un exemple concret une de leurs applications possibles.

Considérons une clinique spécialisée dans le traitement des maladies pulmonaires. Lorsqu'un patient entre dans cette clinique, un médecin doit diagnostiquer une maladie, et donc un traitement, en fonction des symptômes et de l'histoire du patient. Pour l'aider dans sa décision, il pourrait modéliser ses connaissances par un réseau bayésien.

La construction d'un réseau bayésien (figure ci-contre) se fait en quatre étapes détaillées ci-dessous :

1- Création des variables représentant l'univers

On suppose que les trois maladies possibles sont le **cancer** du poumon, la **bronchite** et la **tuberculose**. On suppose qu'on peut faire un examen de **radiographie**, et on considère que les caractéristiques importantes du patient sont son **âge**, son aspect **fumeur** ou non fumeur, un éventuelle **voyage en Asie** et ses éventuels **problèmes respiratoires** chroniques. On introduit de plus une variable **TbOuCa** déterministe pour factoriser le graphe, elle vaut oui s'il le patient a un cancer ou une tuberculose.

2- Définition de l'ensemble des valeurs possibles pour chaque variable

Par exemple, l'âge pourrait prendre trois valeurs : jeune, adulte, senior; et la **bronchite** par oui ou non.

3- Définition des relations de dépendances entre les variables

Par exemple, une éventuelle tuberculose est en relation directe avec un éventuel voyage en Asie.

4- Attribution des probabilités conditionnelles

Par exemple, un patient de la clinique qui a voyagé en Asie a 1% de risque d'avoir la tuberculose.

Si le patient dit qu'il a été en Asie, alors la variable **VisiteAsie** a une probabilité de 100% d'être oui, ensuite on recalcule la nouvelle distribution de probabilités pour les autres variables qui en découlent, ce qui donne des hypothèses plus ou moins probables sur la pathologie réelle du patient.

4 Conclusion

Le théorème de Bayes énoncé au 18^{ème} siècle s'est révélé être un outil très puissant de décision ou d'aide à la décision avec la révolution de l'informatique. Par une approche statistique et probabiliste des données, les réseaux bayésiens permettent d'établir un formalisme des connaissances incertaines, ce qui les place au cœur du domaine l'intelligence artificielle.

Néanmoins, ce formalisme a ses limites. En effet, il modélise l'univers de manière causal et ne permet pas la modélisation de relations d'interdépendances, alors que le besoin s'en fait souvent ressentir. De plus la complexité de la modélisation et des calculs augmente de manière exponentielle par rapport à la taille de l'univers.

Le projet SIGMA2 (Signaux, modèles et Algorithmes) illustre bien les orientations des recherches d'aujourd'hui sur les réseaux bayésiens :

- recherches sur l'identification du comportement dynamique d'un véhicule routier, et sur la surveillance et le diagnostic d'organes de véhicules automobiles, pour les constructeurs d'automobiles.
- recherches sur le diagnostic de pannes dans les réseaux de télécommunications et sur l'identification et la surveillance de structures vibrantes en ambiance de travail.
- recherches sur l'exploitation des données de vol sous excitation naturelle pour les essais en vol.
- recherches sur les chaînes de Markov (traitement des réseaux bayésiens avec données temporelles) et sur les méthodes statistiques pour les systèmes dynamiques stochastiques.

Egalement beaucoup de recherches sont faites pour accélérer les algorithmes et parer la complexité exponentielle des réseaux bayésiens.

5 Bibliographies

[1] : Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, San Mateo, CA.

[2] : Howard, R. A., & Matheson, J. E. (1981). Influence diagrams. In Howard, R. A., & Matheson, J. (Eds.), *The Principles and Applications of Decision Analysis*, page 720–762. Strategic Decisions Group, CA.

[3] : Antoine Cornuéjols, Laurent Miclet, *Apprentissage artificiel: Concepts et algorithmes*, édition Eyrolles, page 364-365.

[4] : Stutz, J. & P. Cheeseman, "A Short Exposition on Bayesian Inference and Probability." June 1994. National Aeronautic and Space Administration Ames Research Centre: Computational Sciences Division, Data Learning Group,

[5] : Thomas Richardson. Bayes Net course, Helsinki, April 1997. Lecture 4 part 2, page 5. <http://www.cs.helsinki.fi/research/cosco/Calendar/BNCCourse/Notes.html>

[6] : Todd A. Stephenson. An introduction to Bayesian network theory and usage. IDIAP-RR 00-03, february 2003