

Réseaux bayésiens

Définition d'un réseaux bayésien

Structure d'un réseaux bayésien

Probabilité dans un RB

Règle de la D-séparation

Inférence dans un réseaux bayésien

Construction d'un réseau bayésien

Les table de probabilités

La structure du réseau bayésien

Applications

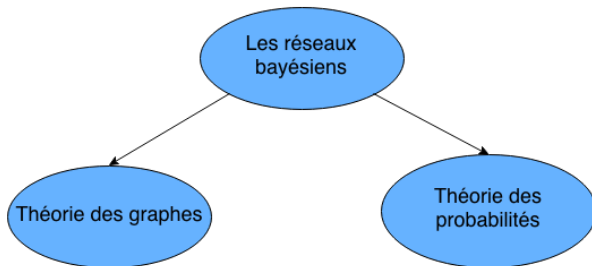
Réseaux bayésiens dynamiques

Chaine de markov de premier ordre

Chaine de markov cachée

Les types d'inférences

- Le raisonnement probabiliste classique n'est pas applicable pour des cas ou des centaines (milliers) de variables aléatoires.
- les RB une façon plus intuitive de le faire



- Représentation des connaissances probabilistes d'une application.

Réseaux
bayésiensDéfinition d'un
réseaux bayésienStructure d'un réseaux
bayésien

Probabilité dans un RB

Règle de la
D-séparationInférence dans un
réseaux bayésienConstruction d'un
réseau bayésien

Les table de probabilités

La structure du réseau
bayésien

Applications

Réseaux
bayésiens
dynamiquesChaîne de markov de
premier ordre

Chaîne de markov cachée

Les types d'inférences

Définition

- C'est un graphe orienté acyclique
 - Les noeuds sont les variables aléatoires
 - les arcs sont les dépendances (causalités) probabilistes entre les variables.
 - C'est la distribution des probabilités conditionnelles

Réseaux
bayésiensDéfinition d'un
réseaux bayésienStructure d'un réseaux
bayésien

Probabilité dans un RB

Règle de la
D-séparationInférence dans un
réseaux bayésienConstruction d'un
réseau bayésien

Les table de probabilités

La structure du réseau
bayésien

Applications

Réseaux
bayésiens
dynamiquesChaîne de markov de
premier ordre

Chaîne de markov cachée

Les types d'inférences

Example

Considérons la situation suivante :

- Je suis au travail et mes voisins Marie et Jean m'appellent chaque fois que mon alarme sonne
- Parfois Jean confond l'alarme avec le téléphone
- Marie met parfois la musique trop fort
- L'alarme peut sonner s'il y a de légers séismes
- Comment conclure qu'il y a un cambriolage chez moi ?

Réseaux bayésiens

Définition d'un réseaux bayésien

Structure d'un réseaux bayésien

Probabilité dans un RB

Règle de la D-séparation

Inférence dans un réseaux bayésien

Construction d'un réseaux bayésien

Les table de probabilités

La structure du réseau bayésien

Applications

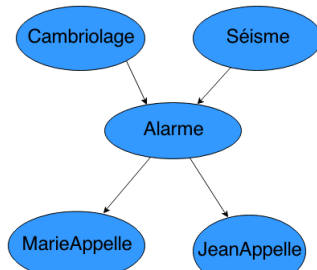
Réseaux bayésiens dynamiques

Chaîne de markov de premier ordre

Chaîne de markov cachée

Les types d'inférences

- La topologie du RB modélise les relations de causalités
 - un cambriolage peut déclencher l'alarme, de même qu'un séisme
 - l'alarme peut inciter Jean à appeler, de même pour Marie
- Un arc d'un noeud X vers un noeud Y signifie que la variable X influence la variable Y
 - X est appelé parent de Y
 - Parents(Y) est l'ensemble des parents de Y



Calcul des probabilités conjointes

- Un réseau bayésien est une façon compacte de représenter des probabilités conjointes.
- La distribution conditionnelle de X_1 sachant X_2 est notée $P(X_1|X_2)$

- $P(X_1|X_2) = \frac{P(X_1, X_2)}{P(X_2)}$

- Soit $X_1 = \{X_1, \dots, X_n\}$, l'ensemble des variables d'un réseau bayésien

$$P(X_1, \dots, X_n) = \prod_{k=1}^n P(X_i | \text{parents}(X_i))$$

Réseaux
bayésiensDéfinition d'un
réseaux bayésienStructure d'un réseaux
bayésien

Probabilité dans un RB

Règle de la
D-séparationInférence dans un
réseaux bayésienConstruction d'un
réseau bayésien

Les table de probabilités

La structure du réseau
bayésien

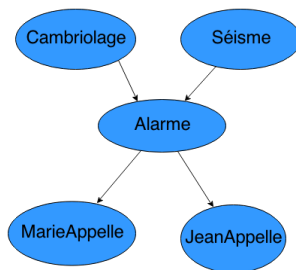
Applications

Réseaux
bayésiens
dynamiquesChaîne de markov de
premier ordre

Chaîne de markov cachée

Les types d'inférences

Calcul des probabilités marginales



$$\begin{aligned}
 P(C = F, A = V) &= \sum_s \sum_j \sum_m P(J = j, M = m, A = V, C = F, S = s) \\
 &= \sum_s \sum_j \sum_m P(j|A = V)P(m|A = V)P(A = V|C = F, s)P(C = F)P(s) \\
 &= \sum_s \sum_j P(j|A = V)P(A = V|C = F, s)P(C = F)P(s) \sum_m P(m|A = V) \\
 &= \sum_s P(A = V|C = F, s)P(C = F)P(s) \sum_j P(j|A = V)
 \end{aligned}$$

$$P(C = F, A = V) = \sum_s P(A = V|C = F, s)P(C = F)P(s)$$

Règle de la D-séparation

- **D-séparation** : critère général pour décider si un noeud X est indépendant d'un noeud Y , étant donné d'autres noeuds $Z = \{Z_1, \dots, Z_m\}$
- X est indépendant de Y sachant Z si tous **les chemins non dirigés** entre X et Y sont **bloqués** par Z
- Un chemin est bloqué s'il contient au moins un noeud N qui satisfait l'une des deux conditions :
 - il inclue un noeud $\rightarrow \textcircled{N} \rightarrow$ ou $\leftarrow \textcircled{N} \leftarrow$ où $N \in Z$
 - il inclue un noeud $\rightarrow \textcircled{N} \leftarrow$ et $N \notin Z$, ni **aucun descendant** de N

Réseaux
bayésiensDéfinition d'un
réseaux bayésienStructure d'un réseaux
bayésien

Probabilité dans un RB

Règle de la
D-séparationInférence dans un
réseaux bayésienConstruction d'un
réseau bayésien

Les table de probabilités

La structure du réseau
bayésien

Applications

Réseaux
bayésiens
dynamiquesChaine de markov de
premier ordre

Chaine de markov cachée

Les types d'inférences

Inférence dans un réseaux bayésien

- C'est un Problème NP complet
- Inférence approximative
 - Méthode de rejet (*Rejection sampling*)
- Résistance aux bruit.

Réseaux
bayésiensDéfinition d'un
réseaux bayésienStructure d'un réseaux
bayésien

Probabilité dans un RB

Règle de la
D-séparationInférence dans un
réseaux bayésien**Construction d'un
réseau bayésien**

Les table de probabilités

La structure du réseau
bayésien

Applications

Réseaux
bayésiens
dynamiquesChaine de markov de
premier ordre

Chaine de markov cachée

Les types d'inférences

Construction d'un réseau bayésien

- Nous avons besoin de deux choses
 - Les tables de probabilités
 - La structure du réseau bayésien

Réseaux
bayésiensDéfinition d'un
réseaux bayésienStructure d'un réseaux
bayésien

Probabilité dans un RB

Règle de la
D-séparationInférence dans un
réseaux bayésienConstruction d'un
réseau bayésien

Les table de probabilités

La structure du réseau
bayésien

Applications

Réseaux
bayésiens
dynamiquesChaine de markov de
premier ordre

Chaine de markov cachée

Les types d'inférences

Les table de probabilités

- On peut faire appel a des experts :
 - fastidieux
 - pas naturel ou intuitif
- Il serait préférable d'automatiser ce process
 - Collecter des données sur l'environnement des exemples statistiques calcule des fréquences
 - On dérive les tables probabilités qui refletent les données
- Si on a un ensemble de données ou certains noeuds ne sont pas observés ou qui'il existe des valeurs manquantes
 - algorithme EM (*Estimation-Maximization*)

Réseaux
bayésiensDéfinition d'un
réseaux bayésienStructure d'un réseaux
bayésien

Probabilité dans un RB

Règle de la
D-séparationInférence dans un
réseaux bayésienConstruction d'un
réseau bayésien

Les table de probabilités

La structure du réseau
bayésien

Applications

Réseaux
bayésiens
dynamiquesChaîne de markov de
premier ordre

Chaîne de markov cachée

Les types d'inférences

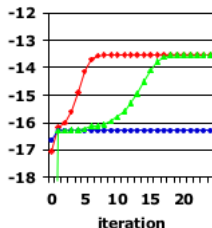
Expectation Maximization

- Distribution initiale
- L'étape E : calculer la fonction de vraisemblance par rapport aux variables non observées.

$$P(H|X, \theta)$$

- L'étape M : Maximiser la nouvelle distribution

$$\theta = \max_{\theta'} \log(P(H|X, \theta'))$$



- Autre algorithme : Gradient Ascent

Réseaux
bayésiensDéfinition d'un
réseaux bayésienStructure d'un réseaux
bayésien

Probabilité dans un RB

Règle de la
D-séparationInférence dans un
réseaux bayésienConstruction d'un
réseau bayésien

Les table de probabilités

La structure du réseau
bayésien

Applications

Réseaux
bayésiens
dynamiquesChaîne de markov de
premier ordre

Chaîne de markov cachée

Les types d'inférences

Algorithme K2

- **Entrée :**
 - Un ensemble de noeuds ordonnées.
 - Une base de données contenant des observations sur ces noeuds.
 - Un nombre max de parents qu'un noeud peut avoir.
- **Sortie :** L'ensemble des parents de chaque noeuds.
- **Fonction à maximiser :**

$$g(i, \pi_i) = \prod_{j=1}^{q_i} \frac{(r_i-1)!}{(N_{ij}+r_i-1)!} \prod_{k=1}^{r_i} N_{ijk}!$$
 - r_i : Une variable X_i à r_i valeur possible
(v_{i1}, \dots, v_{ir_i})
 - N_{ijk} le nombre de cas dans D où X_i à la valeur v_{ik}
 - π_i l'ensemble des parents de X_i

Réseaux
bayésiensDéfinition d'un
réseaux bayésienStructure d'un réseaux
bayésien

Probabilité dans un RB

Règle de la
D-séparationInférence dans un
réseaux bayésienConstruction d'un
réseau bayésien

Les table de probabilités

La structure du réseau
bayésien

Applications

Réseaux
bayésiens
dynamiquesChaine de markov de
premier ordre

Chaine de markov cachée

Les types d'inférences

Algorithme K2

Pour i de 1 à n

$$P = g(i, \pi_i)$$

Tant que $|\pi_i| < u$

- Soit $Z \in Pred(X_i) - \pi_i$ qui maximise $g(i, \pi_i \cup \{Z\})$
- $P_{new} = g(i, \pi_i \cup \{Z\})$
- Si $P_{new} > P_{old}$ alors $P_{old} = P_{new}$ sinon Arrêt

fin tant que

Ecrire(i, π_i)

fin pour

Réseaux
bayésiensDéfinition d'un
réseaux bayésienStructure d'un réseaux
bayésien

Probabilité dans un RB

Règle de la
D-séparationInférence dans un
réseaux bayésienConstruction d'un
réseau bayésien

Les table de probabilités

La structure du réseau
bayésien

Applications

Réseaux
bayésiens
dynamiquesChaîne de markov de
premier ordre

Chaîne de markov cachée

Les types d'inférences

Hill Climbing

- ① on débute avec un graphe acyclique aléatoire comme graphe courant
- ② on obtient ses tables de probabilités à partir des fréquences d'observation du graphe courant
- ③ on utilise la recherche locale pour générer des graphes successeurs du graphe courant
 - i on obtient les tables de probabilités du graphe successeur
 - ii on remplace le graphe courant par le successeur s'il est "meilleur"
- ④ on retourne à 2. jusqu'à un critère d'arrêt

Réseaux bayésiens

Définition d'un réseau bayésien

Structure d'un réseau bayésien

Probabilité dans un RB

Règle de la D-séparation

Inférence dans un réseau bayésien

Construction d'un réseau bayésien

Les table de probabilités

La structure du réseau bayésien

Applications

Réseaux bayésiens dynamiques

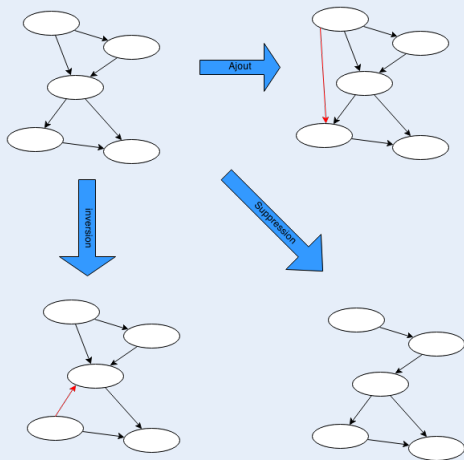
Chaîne de markov de premier ordre

Chaîne de markov cachée

Les types d'inférences

Graphe successeur

- On génère des successeurs à partir des modifications suivantes



Réseaux
bayésiensDéfinition d'un
réseaux bayésienStructure d'un réseaux
bayésien

Probabilité dans un RB

Règle de la
D-séparationInférence dans un
réseaux bayésienConstruction d'un
réseau bayésien

Les table de probabilités

La structure du réseau
bayésien

Applications

Réseaux
bayésiens
dynamiquesChaîne de markov de
premier ordre

Chaîne de markov cachée

Les types d'inférences

Fonction objectif

- La fonction objectif à maximiser par la recherche locale est

$$\sum_t \log P(X_1 = x_1^t, \dots, X_n = x_n^t) - M(\log T)/2$$

- M est le nombre de paramètres requis par les tables de probabilités conditionnelles du réseau
- On cherche un graphe
 - qui explique bien les données (leur donne une haute probabilité)
 - qui est compact (peu de paramètres)

Réseaux bayésiens

Définition d'un réseaux bayésien

Structure d'un réseaux bayésien

Probabilité dans un RB

Règle de la D-séparation

Inférence dans un réseaux bayésien

Construction d'un réseau bayésien

Les table de probabilités

La structure du réseau bayésien

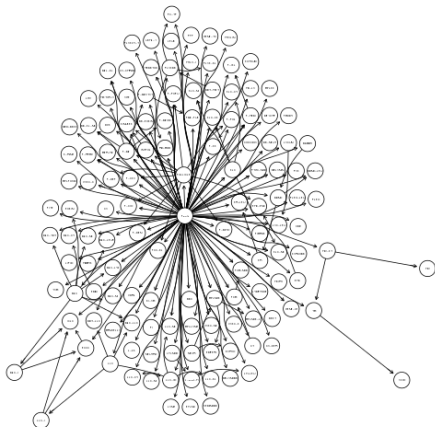
Applications

Réseaux bayésiens dynamiques

Chaine de markov de premier ordre

Chaine de markov cachée

Les types d'inférences



Pathfinder

- Diagnostic des maladies des ganglions lymphatiques
- Nombre de noeuds : 135
- Nombre d'arcs : 200

Réseaux bayésiens

Définition d'un réseaux bayésien

Structure d'un réseaux bayésien

Probabilité dans un RB

Règle de la D-séparation

Inférence dans un réseaux bayésien

Construction d'un réseau bayésien

Les table de probabilités

La structure du réseau bayésien

Applications

Réseaux bayésiens dynamiques

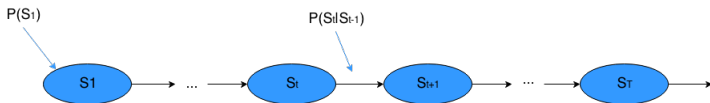
Chaîne de markov de premier ordre

Chaîne de markov cachée

Les types d'inférences

Chaîne de markov de premier ordre

- Une seule serie de variables aléatoires S indicée par le temps t
- On va avoir besoin d'une
 - distribution a priori (initiale)
 - matrice de transition $P(S_{t+1}|S_t)$



- $P(S_{1:T}) = P(S_1) \prod_{t=2}^T P(S_t|S_{t-1})$

Réseaux bayésiens

Définition d'un réseaux bayésien

Structure d'un réseaux bayésien

Probabilité dans un RB

Règle de la D-séparation

Inférence dans un réseaux bayésien

Construction d'un réseau bayésien

Les table de probabilités

La structure du réseau bayésien

Applications

Réseaux bayésiens dynamiques

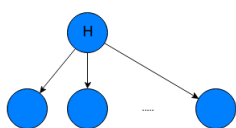
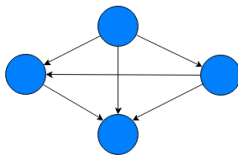
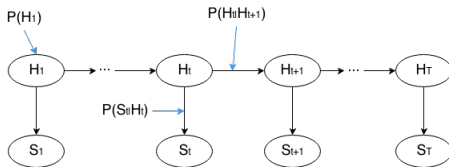
Chaîne de markov de premier ordre

Chaîne de markov cachée

Les types d'inférences

Chaîne de markov cachée

- H_t sont les variables cahcées
- S_t sont les variables d'observation

 $O(n)$ paramètres $O(2^n)$ paramètres

$$P(S_{1:T}, H_{1:T}) = P(H_1|S_1) \prod_{t=2}^T P(H_t|H_{t-1})P(S_t|H_t)$$

Réseaux bayésiens

Définition d'un
réseaux bayésien

Structure d'un réseaux
bayésien

Probabilité dans un RB

Règle de la
D-séparation

Inférence dans un
réseaux bayésien

Construction d'un
réseau bayésien

Les table de probabilités

La structure du réseau
bayésien

Applications

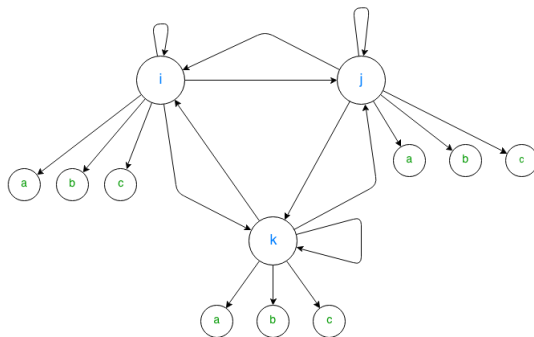
Réseaux
bayésiens
dynamiques

Chaîne de markov de
premier ordre

Chaîne de markov cachée

Les types d'inférences

Chaîne de markov cachée



Probabilité d'une séquence observée

$$P(S_{1:T}) = \sum_{h_{1:T}} P(H_{1:T} = h_{1:T}) P(S_{1:T} | H_{1:T} = h_{1:T})$$

- Un nombre exponentiel de séquence cachées possibles.
- Utiliser la **programmation dynamique**
 - on définit $\alpha(i, t) = P(S_{1:t} = s_{1:t}, H_t = i)$
 - $\alpha(i, t + 1)$ s'écrit en fonction de $\alpha(j, t)$
- une fois le tableau calculé on obtient :

$$P(S_{1:T}) = \sum_j \alpha(j, T)$$

Réseaux
bayésiensDéfinition d'un
réseaux bayésienStructure d'un réseaux
bayésien

Probabilité dans un RB

Règle de la
D-séparationInférence dans un
réseaux bayésienConstruction d'un
réseau bayésien

Les table de probabilités

La structure du réseau
bayésien

Applications

Réseaux
bayésiens
dynamiquesChaîne de markov de
premier ordre

Chaîne de markov cachée

Les types d'inférences

Les types d'inférences

Filtrage

Calcul de la distribution a posteriori de la variable caché la plus récente

- Quelle est la probabilité qu'il pleuve aujourd'hui ?

Prédiction

Calcul de la distribution a posteriori sur un état futur

- Quelle est la probabilité qu'il pleuve dans 3 jours ?

Réseaux
bayésiensDéfinition d'un
réseaux bayésienStructure d'un réseaux
bayésien

Probabilité dans un RB

Règle de la
D-séparationInférence dans un
réseaux bayésienConstruction d'un
réseau bayésien

Les table de probabilités

La structure du réseau
bayésien

Applications

Réseaux
bayésiens
dynamiquesChaine de markov de
premier ordre

Chaine de markov cachée

Les types d'inférences

Les types d'inférences

Lissage

Calcul de la distribution sur un état passé

- Quelle est la probabilité qu'il y ait eu de la pluie hier ?

L'explication la pus plausible

Trouver la séquence d'états cachés qui explique le mieux possible les observations

- Quelle est la météo la plus probable pour toutes les t dernières journées ?

Réseaux bayésiens

Définition d'un réseaux bayésien

Structure d'un réseaux bayésien

Probabilité dans un RB

Règle de la D-séparation

Inférence dans un réseaux bayésien

Construction d'un réseau bayésien

Les table de probabilités

La structure du réseau bayésien

Applications

Réseaux bayésiens dynamiques

Chaîne de markov de premier ordre

Chaîne de markov cachée

Les types d'inférences

Comparaison avec d'autres techniques

Connaissance	Réseaux de neurones	Arbres de décision	Systèmes experts	Réseaux bayésiens
Acquisition				
Expertise seulement			*	+
Données seulement	*	+		*
Mixte	+	+		+
Données incomplètes	+			*
Représentation				
Incertitude			+	*
Lisibilité		+	+	*
Facilité	+	*		
Homogénéité				*
Utilisation				
Requêtes élaborées			+	*
Performances	*			

- + : la technique considérée permet de prendre en compte le problème
- * : la meilleure technique du point de vue de la caractéristique considérée.