TP N° 67

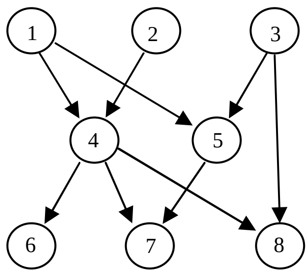
Réseaux bayésiens[1]

L'objet de ce TP est de présenter les réseaux bayésiens et de montrer leur intérêt dans le domaine de la Sûreté de Fonctionnement.

Un réseau bayésien est un modèle graphique de représentation des dépendances entre des variables aléatoires permettant d'effectuer des calculs de probabilité entre ces dernières.

Le graphe peut être construit par un expert ou résulter d'un apprentissage automatique à partir de données observées permettant d'estimer ses paramètres et éventuellement sa structure.

Les réseaux bayésiens sont notamment utilisés pour le diagnostic médical et industriel.



-

1 - Combinaisons d'événements

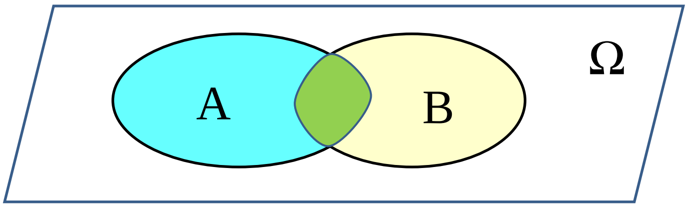
Expliquer les formules de probabilité des combinaisons d'événements, la notion de dépendance, le théorème des probabilités totales et le théorème de Bayes sous sa forme classique et généralisée.

2 - Réseaux bayésiens

Présenter les réseaux bayésiens, les calculs de probabilité associés et des techniques d'apprentissage de leurs paramètres et de leur structure.

[1] Ce TP a été réalisé en collaboration avec Marilyn Renou, stagiaire au CNES Toulouse.

1 - Combinaisons d'événements



Les probabilités des combinaisons entre deux événements A et B dans l'espace des possibles Ω s'expriment de la manière suivante.

Négation : P(non[A]) = $P(\bar{A}) = 1 - p(A)$

Union : $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ Théorème de Poincaré

Probabilité conditionnelle : $P(A/B)$ = P(A sachant B) = $P(A \cap B) / P(B)$ avec $P(B) \neq 0$

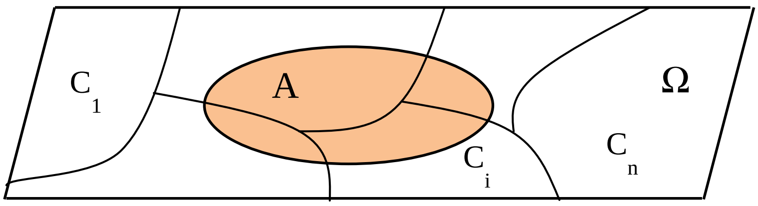
Intersection : $P(A \cap B) = P(A/B)\, P(B) = P(B/A)\, P(A)$

Evénements indépendants : $P(A/B) = P(A)$ soit $P(A \cap B) / P(B) = P(A) \Leftrightarrow P(A \cap B) = P(A)\, P(B)$

$\Rightarrow P(A \cap B) = P(A)\, P(B)$ si et seulement si les événements A et B sont indépendants

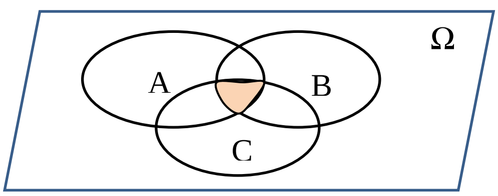
Théorème de Bayes : $P(A \cap B) = P(A/B)\, P(B) = P(B/A)\, P(A) \Rightarrow P(A/B) = P(B/A)\, P(A) / P(B)$

Théorème des probabilités totales : Permet de calculer la probabilité d'un événement A en considérant différentes conditions disjointes C_i recouvrant l'ensemble Ω.



$P(A) = \sum P(A \cap C_i) = \sum P(A/C_i) \times P(C_i)$ avec $P(C_i \cap C_j) = 0$ et $\sum P(C_i) = 1$

Théorème de Bayes généralisé : Permet de calculer la loi de probabilité jointe à plusieurs événements (intersection).

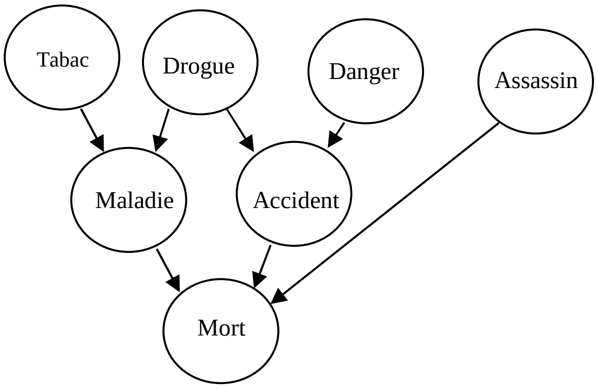


$P(A, B) = P(A \cap B) = P(A)\, P(B/A)$ $P(A, B, C) = P(A \cap B \cap C) = P(A)\, P(B/A)\, P(C/A,B)$

Soit en généralisant : $P(X_1 \ldots X_n) = P(X_1)\, P(X_2 / X_1) \ldots P(X_n / X_1 \ldots X_{n-1}) = \Pi_1^n P(X_i / X_1 \ldots X_{i-1})$

2 - Réseaux bayésiens

Un réseau bayésien représente les liens de dépendance entre des variables aléatoires sous forme graphique, comme dans l'exemple ci-dessous.



Les variables aléatoires peuvent être de type booléen (vrai, faux dans cet exemple), discret, voire continu.

Les variables aléatoires sont représentées par des cercles et les liens de dépendance par des flèches ou arcs orientés dont le sens représente un lien de parenté (parent → enfant) généralement causal (cause → effet) quand le graphe est construit par un expert (et non pas généré automatiquement à partir des données).

Le graphe orienté est acyclique (sans boucle), puisqu'une variable ne peut pas dépendre indirectement d'elle-même, et les parents de chacune des variables sont considérés indépendants entre eux tant que l'état de leur enfant n'est pas connu.

A partir d'une observation (ensemble de variables instantiées), le réseau bayésien permet de calculer la probabilité de chacun des événements. Ainsi est-il possible d'estimer, dans cet exemple, le risque de mortalité d'un fumeur en présence d'un assassin ou inversement d'estimer la probabilité d'assassinat en présence d'un mort.

Différents types de connexions apparaissent dans le graphe :

- Série : Tabac → Maladie → Mort
- Divergente (structure en Λ) : Maladie ← Drogue → Accident
- Convergente (structure en V) : Maladie → Mort ← Accident

Dans la première connexion, le Tabac et la Mort sont dits **indépendants conditionnellement** à la Maladie car le Tabac n'apporte pas d'information à la mort si la Maladie est connue.

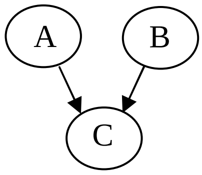
Dans la seconde connexion, la Maladie et l'Accident sont dits **indépendants conditionnellement** à la Drogue car la Maladie n'apporte pas d'information à l'Accident (ou inversement) si la Drogue est connue.

Dans la troisième connexion, la Maladie et l'Accident sont dits **dépendants conditionnellement** à la Mort car la Maladie apporte de l'information à l'Accident (ou inversement) si la Mort est connue (s'il n'est pas mort d'accident c'est probablement de maladie).

Deux variables quelconques sont dites indépendantes conditionnellement à un ensemble de variables connues (d-séparées) si tous les chemins qui les relient passent par une connexion :

- de type série ou divergente dont la variable intrermédiaire est connue
- ou de type convergente dont la variable intrermédiaire et les descendants sont inconnus.

Quand l’état d’un enfant et de sa descendance est inconnu, sa probabilité peut s’exprimer à partir de celle de ses parents (considérés indépendants) par application du théorème des probabilités totales :



$$P(C) = P(A)\ P(B)\ P(C/AB) + P(A)\ P(\cancel{B})\ P(C/A\cancel{B}) + P(\cancel{A})\ P(B)\ P(C/\cancel{A}B) + P(\cancel{A})\ P(\cancel{B})\ P(C/\cancel{A}\cancel{B})$$

$$\text{Avec } P(\cancel{X}) = 1 - P(X)$$

Si l’état de l’enfant est connu, l’expression de la loi jointe aux divers événements (calculée par application du théorème de Bayes généralisé) se simplifie en raison de l’indépendance des parents.

$$P(A, B, C) = P(A)\ P(B/A)\ P(C/A, B) = P(A)\ P(B)\ P(C/A, B)$$

Appliqué à tout ensemble d’événements du graphe, la loi jointe globale se décompose en un produit de lois conditionnelles locales :

$$P(x_1...x_n) = \Pi_1^n\ P(x_i / x_1…x_{i-1}) = \Pi_1^n\ P(x_i / \text{parents}(x_i))$$

La probabilité conditionnelle de A sachant C peut se calculer de la manière suivante par application du théorème des probabilités totales :

$$P(A/C \text{ vrai}) = P(A, B, C) / P(C) = [P(A)\ P(B)\ P(C/A,B) + P(A)\ P(\cancel{B})\ P(C/A,\cancel{B})] / P(C)$$

ou

$$P(A/C \text{ faux}) = P(A, B, \cancel{C}) / [1 - P(C)] = [P(A)\ P(B)\ P(\cancel{C}/A,B) + P(A)\ P(\cancel{B})\ P(\cancel{C}/A,\cancel{B})] / [1 - P(C)]$$

Si l’état d’un enfant n’est pas connu mais seulement l’un de ses descendants, la probabilité conditionnelle de A devient :

$$P(A/ \text{observation}) = P(C/\text{observation})\ [P(A)\ P(B)\ P(C/A,B) + P(A)\ P(\cancel{B})\ P(C/A,\cancel{B})] / P(C)$$
$$+\ [1- P(C/\text{observation}]\ [P(A)\ P(B)\ P(\cancel{C}/A,B) + P(A)\ P(\cancel{B})\ P(\cancel{C}/A,\cancel{B})] / [1 - P(C)]$$

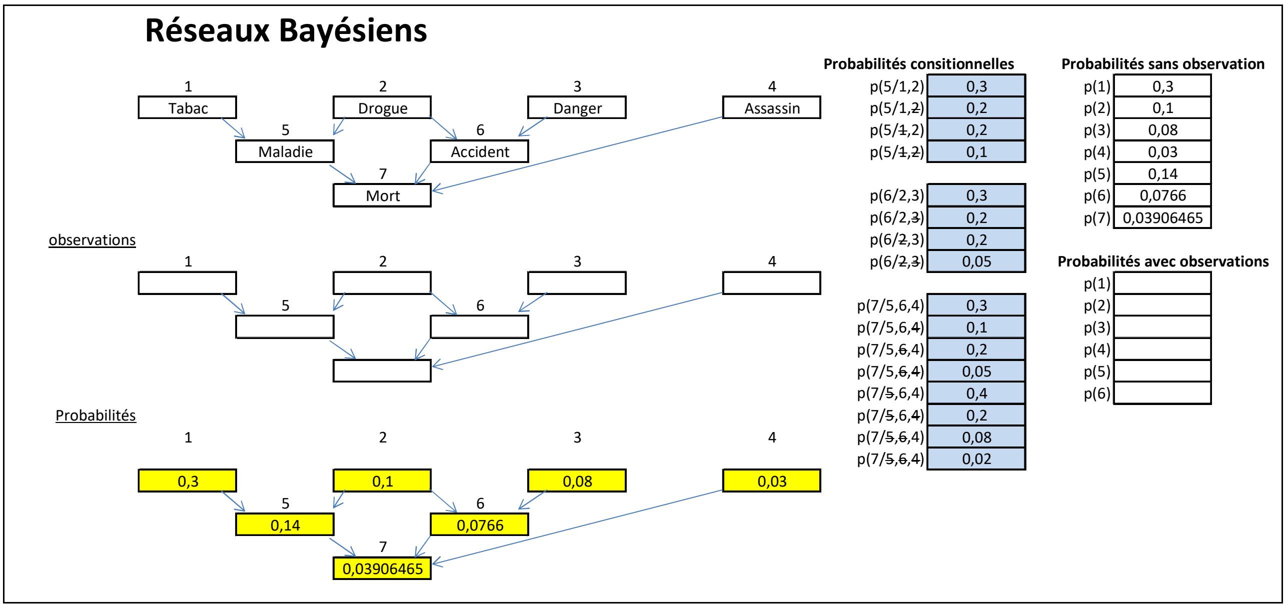
Ainsi, la probabilité de chacun des événements du graphe peut se calculer de proche en proche en partant des événements sans descendants connus à l’exeption de ses propres enfants.

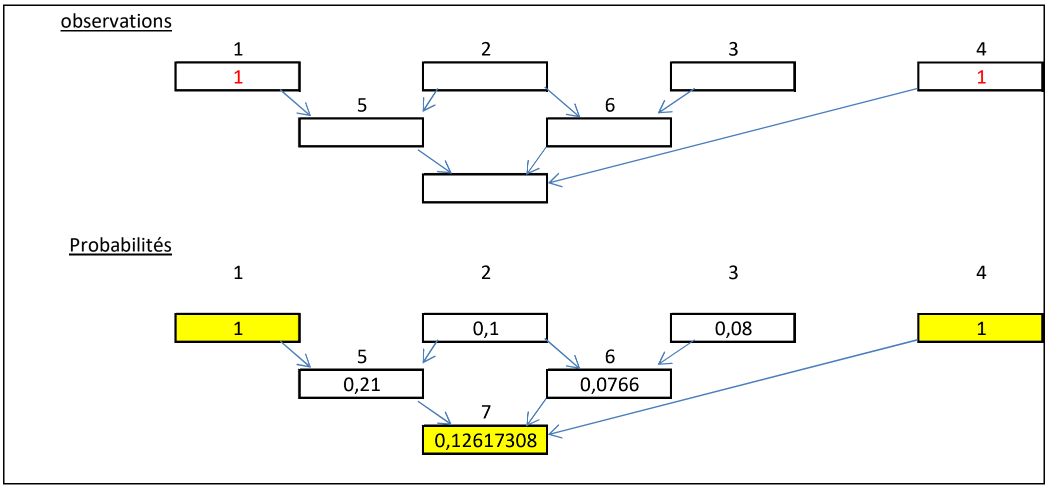
L’exemple précédent traité ci-après sous Excel estime le risque de mortalité sans élément observé, celui d’un fumeur en présence d’un assassin et la probabilité d’assassinat en présence d’un mort. On notera dans cet exemple que les causes de mortalité ne sont pas disjointes car un malade est considéré moins résistant en cas d’accident, par exemple.

L’apprentissage automatique consiste à extraire de l’information contenue dans des données pour estimer les paramètres d’un réseau bayésien, voire définir la structure même du graphe.

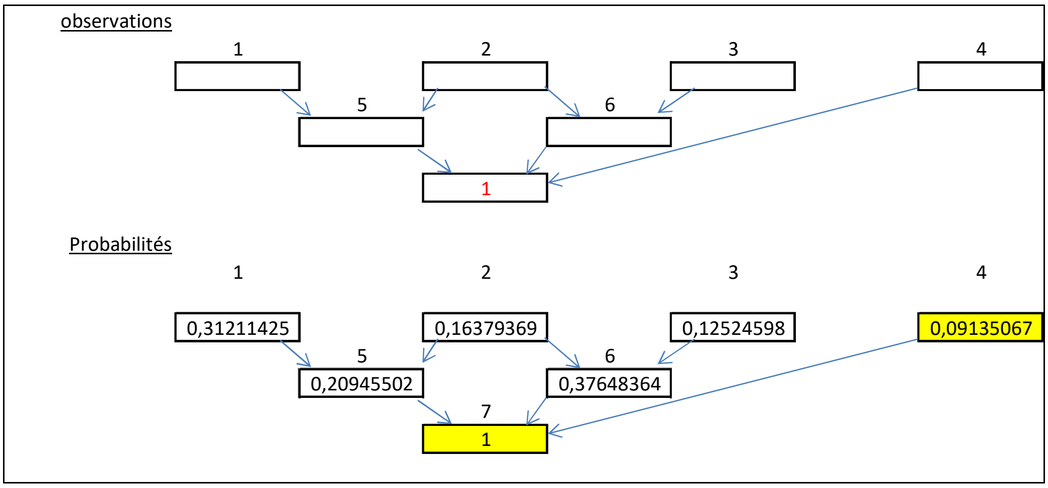
Le calcul de la fréquence d’apparition des événements sous certaines conditions permet d’estimer les probabilités conditionnelles ou d’enrichir leur connaissance a priori.

Fondée sur des techniques d’intelligence artificielle, la construction du graphe se base sur l’analyse des corrélations entre les variables pour ajouter ou supprimer des arêtes et détecter divers types de structure. Le graphe généré n’a alors aucune raison d’être de type causal





Mortalité d’un fumeur en présence d’un assassin



Probabilité d’assassinat en présence d’un mort

Feuille de calcul
Microsoft Excel