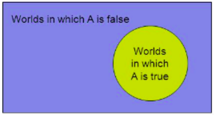
Rețele bayesiene

IA 2023/2024

Conținut

Probabilitatea asociată unei propoziții este suma probabilităților lumilor în care aceasta este adevărată.

$$P(\phi) = \sum_{w \in \phi} P(w), \quad \phi \ \textit{propoziție}$$



P(A) is the area of the oval

Exemplu: $P(Total = 11) = P((5,6)) + P((6,5)) = 1/36 + 1/36 = 1/18$.

Presupunem $P(doubles) = 1/4$.

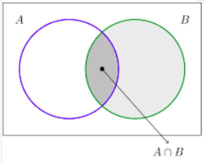
Probabilități necondiționate (*prior*)

Probabilități condiționate

Probabilități condiționate (*posterior*) $P(A|B)$ este fracțiunea de lumi posibile în care B este adevărată și atunci și A este adevărată

- probabilitatea lui A, dat fiind B

$$P(a|b) = \frac{P(a \wedge b)}{P(b)}$$



Exemplu: $P(doubles|Die_1 = 5) = \frac{P(doubles \wedge Die_1 = 5)}{P(Die_1 = 5)}$.

Regula produs: $P(a \wedge b) = P(a|b)P(b)$

Probabilități

- Distribuție de probabilitate a unei variabile aleatoare
  $P(Weather = sunny) = 0.6$
  $P(Weather = rain) = 0.1$
  $P(Weather = cloudy) = 0.29$
  $P(Weather = snow) = 0.01$

  $\mathbf{P}(Weather) = \langle 0.6, 0.1, 0.29, 0.01 \rangle$

- Funcția densitate de probabilitate pentru o var. aleatoare continuă
  $P(NoonTemp = x) = Uniform_{[18C,26C]}(x)$ temperatura la prânz e distribuită uniform între 18C și 26C

$$P(x) = lim_{dX \to 0} P(x \leq X \leq x + dx)/dx$$

Distribuția comună de probabilitate

$\mathbf{P}(Weather, Cavity)$: probabilitățile tuturor combinațiilor de valori ale var. $Weather$ și $Cavity$. (o tabela 4 x 2)
Obs: Putem utiliza $\mathbf{P}(Weather, Cavity) = \mathbf{P}(Weather|Cavity)\mathbf{P}(Cavity)$ în locul celor 8 ecuații.

Full joint probability distribution: distribuția comună a tuturor variabilelor aleatore
$Cavity, Toothache, Weather \rightarrow \mathbf{P}(Cavity, Toothache, Weather)$

Conținut

Inferență probabilistă: calculul probabilităților condiționate, date fiind anumite observații.

Exemplu:
Distribuția comună de probabilitate:

	toothache		¬toothache	
	catch	¬catch	catch	¬catch
cavity	0.108	0.012	0.072	0.008
¬cavity	0.016	0.064	0.144	0.576

Variabilele bool. *Toothache*, *Cavity*, *Catch*; 3 variabile binare: $2^3-1=7$ parametri independenți;
Pentru n variabile booleene, tabela are dimensiunea $O(2^n)$.

$$P(cavity \vee toothache) = 0.108 + 0.012 + 0.072 + 0.008 + 0.016 + 0.064 = 0.28$$

Distribuție marginală

Distribuția marginală, peste o submulțime de var. (o var.)

- Marginalizare: însumăm probabilitățile pentru fiecare valoare posibilă a celorlalte variabile.

$$\mathbf{P}(\mathbf{Y}) = \sum_{\mathbf{z} \in \mathbf{Z}} \mathbf{P}(\mathbf{Y}, \mathbf{z})$$

Exemplu:

	toothache		¬toothache	
	catch	¬catch	catch	¬catch
cavity	0.108	0.012	0.072	0.008
¬cavity	0.016	0.064	0.144	0.576

$\mathbf{P}(Cavity) = \sum_{\mathbf{z} \in \{Catch, Toothache\}} \mathbf{P}(Cavity, \mathbf{z})$
Prob. marginală $P(cavity) = 0.108 + 0.012 + 0.072 + 0.008 = 0.2$

- Condiționare $\mathbf{P}(\mathbf{Y}) = \sum_{\mathbf{z}} \mathbf{P}(\mathbf{Y}|\mathbf{z}) P(\mathbf{z})$

Probabilități condiționate

Exemplu:

	toothache		¬toothache	
	catch	¬catch	catch	¬catch
cavity	0.108	0.012	0.072	0.008
¬cavity	0.016	0.064	0.144	0.576

$$P(cavity|toothache) = \frac{P(cavity \wedge toothache)}{P(toothache)} = \frac{0.108 + 0.012}{0.108 + 0.012 + 0.016 + 0.064} = 0.6$$

$$P(\neg cavity|toothache) = \frac{P(\neg cavity \wedge toothache)}{P(toothache)} = \frac{0.016 + 0.064}{0.108 + 0.012 + 0.016 + 0.064} = 0.4$$

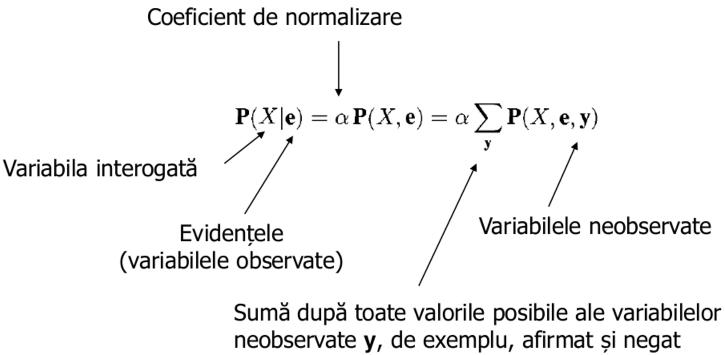
$$\begin{aligned}
\mathbf{P}(Cavity|toothache) &= \alpha \mathbf{P}(Cavity, toothache) \\
&= \alpha[\mathbf{P}(Cavity, toothache, catch) + \mathbf{P}(Cavity, toothache, \neg catch)] \\
&= \alpha[\langle 0.108, 0.016\rangle + \langle 0.012, 0.064\rangle] = \alpha\langle 0.12, 0.08\rangle = \langle 0.6, 0.4\rangle
\end{aligned}$$

Obs: termenul $1/P(toothache)$ const. - const. de normalizare pentru distribuția $\mathbf{P}(Cavity|toothche)$.

Inferență

Procedura de inferență: fie variabila X (*Cavity*), $\mathbf{E}$ lista de variabile evidență (*Toothache*), $\mathbf{e}$ lista de valori observate, $\mathbf{Y}$ variabile neobservate (*Catch*)

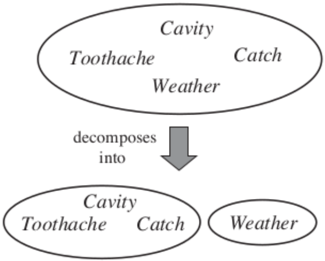
$$\mathbf{P}(X|\mathbf{e}) = \alpha \mathbf{P}(X, \mathbf{e}) = \alpha \sum_{\mathbf{y}} \mathbf{P}(X, \mathbf{e}, \mathbf{y}) \tag{1}$$



Conținut

Independență

Variabilele X și Y sunt independente: $\mathbf{P}(X|Y) = \mathbf{P}(X)$ sau $\mathbf{P}(Y|X) = \mathbf{P}(Y)$ sau $\mathbf{P}(X, Y) = \mathbf{P}(X)\mathbf{P}(Y)$.



Mai puține informații necesare pentru a specifica distribuția comună de probabilitate. Distribuția comună poate fi *factorizată* în două distribuții.

Independență

Exemplu: adăugăm variabila *Weather*.

- $P(toothache, catch, cavity, cloudy) =$
  $P(cloudy|toothache, catch, cavity)P(toothache, catch, cavity)$
  (utilizăm regula produs)

- Obs: problemele dentare nu influențează vremea și invers
  $P(cloudy|toothache, catch, cavity) = P(cloudy)$
  independență (marginală, absolută)

- Deducem
  $P(toothache, catch, cavity, cloudy) = P(cloudy)P(toothache, catch, cavity)$
  și
  $\mathbf{P}(Toothache, Catch, Cavity, Weather) = \mathbf{P}(Toothache, Catch, Cavity)\mathbf{P}(Weather)$

Conținut

Teorema lui Bayes, '73

Utilizând regula produs:

$$P(a \wedge b) = P(a|b)P(b)$$
$$P(a \wedge b) = P(b|a)P(a)$$

$\rightarrow$

$$P(b|a) = \frac{P(a|b)P(b)}{P(a)}$$

$$\mathbf{P}(Y|X) = \frac{\mathbf{P}(X|Y)\mathbf{P}(Y)}{\mathbf{P}(X)}$$

Teorema lui Bayes

$$\mathbf{P}(I|E) = \frac{\mathbf{P}(E|I)\mathbf{P}(I)}{\mathbf{P}(E)}$$

- I ipoteza, E evidența (provine din datele observate)
- $\mathbf{P}(E|I)$ verosimilitatea (*likelihood*) (măsura în care s-a observat evidența, în condițiile îndeplinirii ipotezei)
- $\mathbf{P}(I)$ probabilitatea a priori (*prior*) (gradul de încredere în ipoteză)
- $\mathbf{P}(I|E)$ probabilitatea a posteriori a ipotezei, dată fiind evidența (*posterior*)

Teorema lui Bayes

Cunoaștem evidența (efectul unei cauze necunoscute), și dorim să determinăm cauza:

$$P(cause|effect) = \frac{P(effect|cause)P(cause)}{P(effect)}$$

Diagnostic medical: doctorul cunoaște $P(symptoms|disease)$ și identifică diagnosticul $P(disease|symptoms)$.

Exemplu: $P(s|m) = 0.7$ meningita cauzează înțepenirea gâtului în 70% din cazuri, $P(m) = 1/50000$, $P(s) = 0.01$.

$$P(m|s) = \frac{P(s|m)P(m)}{P(s)} = \frac{0.7 \cdot 1/50000}{0.01} = 0.0014$$

Teorema lui Bayes

Forma generală

$$\mathbf{P}(Y|X) = \alpha \mathbf{P}(X|Y)\mathbf{P}(Y)$$

α const de normalizare

Teorema lui Bayes

Dacă avem mai mult de o variabilă evidență?

- Dacă cunoaștem distribuția comună de probabilitate $P(Cavity|toothache \wedge catch) = \alpha\langle 0.108, 0.016\rangle \approx \langle 0.871, 0.129\rangle$.
- Utilizând teorema lui Bayes:

$$\begin{aligned}\mathbf{P}(Cavity|toothache \wedge catch) &= \alpha\mathbf{P}(toothache \wedge catch|Cavity)\mathbf{P}(Cavity) \\ &= \alpha P(toothache|Cavity)P(catch|Cavity)P(Cavity)\end{aligned}$$

  - $\mathbf{P}(toothache \wedge catch|Cavity) = \mathbf{P}(toothache|Cavity)\mathbf{P}(catch|Cavity)$ Independență condiționată a *toothache* și *catch*, dat *Cavity* (independente condiționat, dată prezența/absența cariei)

Teorema lui Bayes

Independență condiționată a două variabile X și Y, dat Z:
$\mathbf{P}(X, Y|Z) = \mathbf{P}(X|Z)\mathbf{P}(Y|Z)$.

Exemplu:

$$\mathbf{P}(Toothache, Catch|Cavity) = \mathbf{P}(Toothache|Cavity)\mathbf{P}(Catch|Cavity)$$

$$\begin{aligned}
&\mathbf{P}(Toothache, Catch, Cavity) \\
&\quad = \mathbf{P}(Toothache, Catch|Cavity)\mathbf{P}(Cavity) \\
&\quad = \mathbf{P}(Toothache|Cavity)\mathbf{P}(Catch|Cavity)\mathbf{P}(Cavity)
\end{aligned}$$

Pentru n simptome indep cond, dat *Cavity*, dimensiunea reprezentării crește liniar.

$\mathbf{P}(Cause, Effect_1, \ldots, Effect_n) = \mathbf{P}(Cause) \prod_i \mathbf{P}(Effect_i|Cause)$
(modelul *Naive Bayes*)

Independență și independență condiționată

Exemplul 1. Ion (A) și Maria (B) dau cu banul de 100 de ori. Fiecare are un ban diferit.

- evenimente independente
  $P(A|B) = P(A), P(B|A) = P(B)$
  Rezultatul unui experiment nu influențează rezultatul celuilalt experiment

Independență și independență condiționată

Exemplul 2. Ion și Maria dau cu același ban

- există posibilitatea ca banul să nu fie corect
- dacă banul nu este corect, evenimentul A (Ion) poate aduce informații asupra evenimentului B (Maria)
- evenimentele nu sunt independente

Independență și independență condiționată

Exemplul 3. Ion și Maria locuiesc în zone diferite ale orașului și vin la serviciu cu tramvaiul, respectiv mașina

- "Ion a întârziat" și "Maria a întârziat" pot fi considerate independente
- dacă vatmanii sunt în grevă, atunci și traficul rutier crește; evenimentele sunt independente condiționat

Conținut

Rețele bayesiene

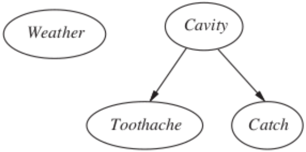
- Modele grafice probabilistice
- Reprezintă dependențele între variabile
- Reprezentarea informațiilor legate de evenimente probabilistice
  - ne va ajuta să realizăm eficient raționamente
- Aplicații: sisteme pentru diagnostic medical, estimarea caracteristicilor psihologice din teste, etc.

Rețea bayesiene

Este un digraf aciclic

- fiecare nod corespunde unei variabile aleatoare (eveniment)
- un arc de la X la Y: X este părintele lui Y (relație: X are o influență directă asupra lui Y)
- fiecare nod X_i are o distribuție de probabilitate condiționată $\mathbf{P}(X_i|Parents(X_i))$ (efectul părinților asupra nodului)

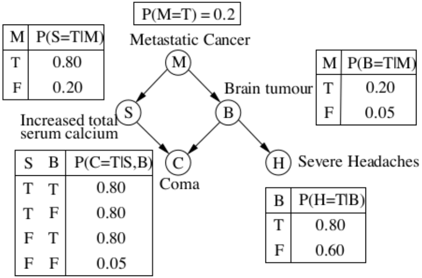
Topologia rețelei specifică relațiile de independență condiționată:



Toothache și *Catch* sunt independente condiționat, dat *Cavity*.

Rețele bayesiene: exemplu

Cancerul metastatic este o cauză posibilă a tumorilor cerebrale și este de asemenea, o explicație pentru creșterea calciului seric total. Oricare dintre acestea ar putea explica intrarea unui pacient în comă. Cefaleea severă este, de asemenea, asociată cu tumorile cerebrale.



M	P(S=T|M)
T	0.80
F	0.20

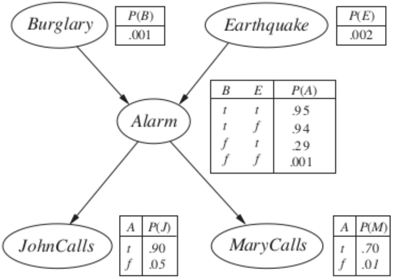
M	P(B=T|M)
T	0.20
F	0.05

S	B	P(C=T|S,B)
T	T	0.80
T	F	0.80
F	T	0.80
F	F	0.05

B	P(H=T|B)
T	0.80
F	0.60

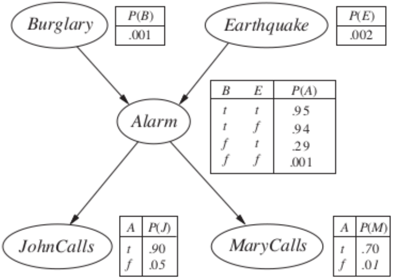
Rețele bayesiene: exemplu

Un sistem de alarmă care sună în cazul unei spargeri, dar și în cazul unui cutremur. Vecinii John și Mary îl sună pe proprietar la serviciu dacă aud alarma.



Efracțiile și cutremurele afectează probabilitatea declanșării alarmei. Alarma influențează probabilitatea ca John și Mary să sune.

Rețele bayesiene: exemplu



Alaturat avem tabelele de probabilitate condiționată.
10 parametri independenți vs. 31.

Dorim să estimăm probabilitatea unei spargeri, în funcție de cine a sunat.

Interogări simple

O conjuncție $P(X_1 = x_1 \wedge \ldots \wedge X_n = x_n)$

Presupunerea modelului bazat pe rețele bayesiene este că o variabilă nu depinde decât de părinții săi:

$$P(x_1, \ldots x_n) = \prod_{i=1}^{n} P(x_i | parents(X_i)) \quad (2)$$

Exemplu: probabilitatea declanșării alarmei, când nu a fost o spargere sau un cutremur, iar John și Mary au sunat

$$\begin{aligned} P(j, m, a, \neg b, \neg e) &= P(j|a)P(m|a)P(a|\neg b \wedge \neg e)P(\neg b)P(\neg e) \\ &= 0.90 \cdot 0.70 \cdot 0.001 \cdot 0.999 \cdot 0.998 = 0.000628 \end{aligned}$$

Construcția unei rețele bayesiene

Utilizăm regula produs pentru a rescrie distribuția comună de probabilitate:

$$
\begin{aligned}
P(x_1, \ldots, x_n) &= P(x_n | x_{n-1}, \ldots, x_1) P(x_{n-1}, \ldots, x_1) \\
&= P(x_n | x_{n-1}, \ldots, x_1) P(x_{n-1} | x_{n-2}, \ldots, x_1) \ldots P(x_2 | x_1) P(x_1) \\
&= \prod_{i=1}^{n} P(x_i | x_{i-1}, \ldots, x_1)
\end{aligned}
$$

(**Regula de înmulțire a probabilităților (chain rule)**)

$$\mathbf{P}(X_i | X_{i-1}, \ldots, X_1) = \mathbf{P}(X_i | Parents(X_i)) \qquad (3)$$

$\forall X_i$ variabilă din rețea, cu condiția $Parents(X_i) \subseteq \{X_{i-1}, \ldots, X_1\}$.

Rețeaua bayesiană este o reprezentare corectă dacă fiecare nod este independent condiționat de predecesorii din ordonare, dați fiind părinții.

Construcția unei rețele bayesiene

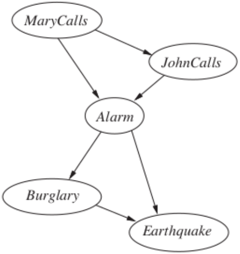
- Determină mulțimea de variabile $\{X_1, \ldots, X_n\}$. Ordonează variabilele a.i. cauzele preced efectele.
- Pentru $i = 1, \ldots n$
  - alege din $X_1, \ldots, X_{i-1}$ o mulțime minimală de părinți a.i. ecuația (3) este satisfacută
  - pentru fiecare părinte, inserează un arc de la acesta la X_i
  - adaugă tabela de probabilitate condiționată $\mathbf{P}(X_i|Parents(X_i))$

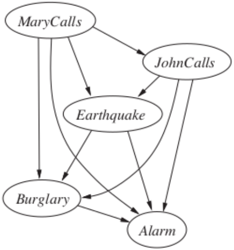
Exemplu:
$P(MaryCalls|JohnCalls, Alarm, Earthquake, Burglary) = P(MaryCalls|Alarm)$,
rezultă *Alarm* este singurul părinte al *MaryCalls*

Rețele bayesiene

- n variabile, fiecare influențată de cel mult k variabile $\rightarrow 2^k$ (pentru a specifica o tabelă de probabilitate condiționată) $\rightarrow n2^k$
  vs.
  Distribuția comună: 2^n
  Exemplu: n=30 noduri, $k = 5$ părinți $\rightarrow$ 960 vs. 10^9.

- Structura rețelei depinde de ordonarea nodurilor





Sortarea topologică

Sortarea topologică a unui digraf este o ordonare liniară a nodurilor a.i. $\forall A \to B$, A apare înaintea lui B

- pentru o rețea bayesiană, sortarea topologică asigură faptul că părinții vor apărea înaintea fiilor
- dacă graful este orientat aciclic, există cel puțin o soluție; dacă există cicluri, sortarea topologică nu este posibilă

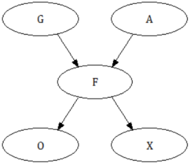
Sortare topologică - algoritmul lui Kahn

```
L ← Empty list that will contain the sorted elements
S ← Set of all nodes with no incoming edge

while S is not empty do
    remove a node n from S
    add n to L
    for each node m with an edge e from n to m do
        remove edge e from the graph
        if m has no other incoming edges then
            insert m into S

if graph has edges then
    return error   (graph has at least one cycle)
else
    return L   (a topologically sorted order)
```

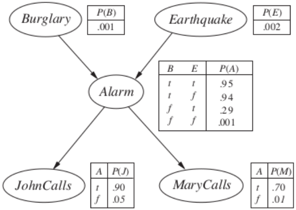
Complexitate timp: $O(n+m)$, n noduri, m arce



1. $L=\emptyset, S=\{G,A\}$
2. $L=\{G\}, S=\{A\}$
3. elimină (GF)
   F nu poate fi adăugat în S: $\exists(AF)$
4. $L=\{G,A\}, S=\emptyset$
5. elimină (AF), $S=\{F\}$
6. $L=\{G,A,F\}, S=\emptyset$
7. elimină (FO), $S=\{O\}$, ...
   $\rightarrow L=\{G,A,F,O,X\}$

Independență condiționată

- Fiecare variabilă este independentă condiționat de ne-descendenți, dați părinții.



JohnCalls este independent de *Burglary*, *Earthquake*, *MarryCalls*, dat *Alarm*.

- *Markov blanket*: un nod e independent condiționat de celălalte noduri, dați părinții, copiii și părinții copiilor

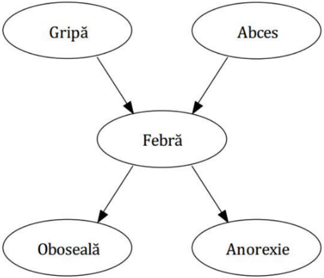
Burglary este independent de *JohnCalls* și *MaryCalls*, dat *Alarm* și *Earthquake*

Conținut

Inferența probabilităților marginale

- Calculează probabilitățile nodurilor, în lipsa unor noduri evidență
- Pentru un nod, calculăm suma probabilităților condiționate de combinațiile posibile de valori ale părinților, înmulțite cu probabilitățile părinților de a avea valorile respective.

Inferența probabilităților marginale: exemplu



$P(Gripă = Da)$	$P(Gripă = Nu)$
0,1	0,9

$P(Abces = Da)$	$P(Abces = Nu)$
0,05	0,95

Gripă	*Abces*	$P(Febră = Da)$	$P(Febră = Nu)$
Da	Da	0,8	0,2
Da	Nu	0,7	0,3
Nu	Da	0,25	0,75
Nu	Nu	0,05	0,95

Febră	$P(Oboseală = Da)$	$P(Oboseală = Nu)$
Da	0,6	0,4
Nu	0,2	0,8

Febră	$P(Anorexie = Da)$	$P(Anorexie = Nu)$
Da	0,5	0,5
Nu	0,1	0,9

Exemplu: Febră

$$
\begin{aligned}
P(f) &= P(f|g,a)P(g)P(a) + P(f|g,\neg a)P(g)P(\neg a) \\
&\quad + P(f|\neg g,a)P(\neg g)P(a) + P(f|\neg g,\neg a)P(\neg g)P(\neg a) \\
&= 0.8 \cdot 0.1 \cdot 0.05 + 0.7 \cdot 0.1 \cdot 0.95 + 0.25 \cdot 0.9 \cdot 0.05 + 0.05 \cdot 0.9 \cdot 0.95 \\
&= 0.1245 \\
P(\neg f) &= 1 - P(f) = 0.8755
\end{aligned}
$$

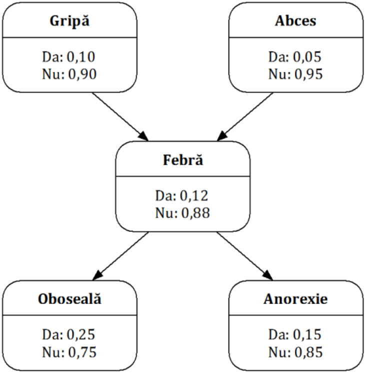
Obs: într-o rețea bayesiană, un nod poate avea oricâte valori posibile.

Exemplu: Oboseală și Anorexie

$$
\begin{aligned}
P(o) &= P(o|f)P(f) + P(o|\neg f)P(\neg f) \\
&= 0.6 \cdot 0.1245 + 0.2 \cdot 0.8755 = 0.2498 \\
P(\neg o) &= 1 - P(o) = 0.7502
\end{aligned}
$$

$$
\begin{aligned}
P(x) &= P(x|f)P(f) + P(x|\neg f)P(\neg f) \\
&= 0.5 \cdot 0.1245 + 0.1 \cdot 0.8755 = 0.1498 \\
P(\neg x) &= 1 - P(x) = 0.8502
\end{aligned}
$$

Exemplu: probabilitățile marginale ale nodurilor



Inferență în rețele bayesiene

- Sistem de inferență probabilist: calculăm distribuția de probabilitate aposteriori pentru o mulțime de variabile, dat un eveniment
- X variabila, $\mathbf{E}$ mulțimea de variabile evidență $E_1, \ldots, E_m$, $\mathbf{e}$ eveniment, $\mathbf{Y}$ non-evidența, $Y_1, \ldots Y_l$ variabile ascunse

  Distribuția de probabilitate aposteriori $\mathbf{P}(X|\mathbf{e}) = ?$

  Exemplu: observăm evenimentul în care $JohnCalls = true$ și $MaryCalls = true$; atunci probabilitatea unei efracții: $\mathbf{P}(Burglary|JohnCalls = true, MaryCalls = true) = \langle 0.284, 0.716 \rangle$.

Inferență prin enumerare

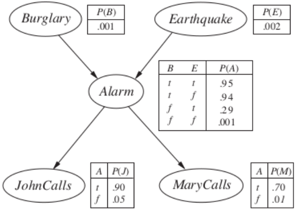
Probabilitatea condiționată

$$\mathbf{P}(X|\mathbf{e}) = \alpha \mathbf{P}(X, \mathbf{e}) = \alpha \sum_{\mathbf{y}} \mathbf{P}(X, \mathbf{e}, \mathbf{y}) \tag{1}$$

O rețea bayesiană oferă o reprezentare a distribuției comune.
Conform ecuației (2), termenul $\mathbf{P}(X, \mathbf{e}, \mathbf{y})$ poate fi scris ca produs de probabilități condiționate.

Inferență prin enumerare: exemplu

Interogare: Care este probabilitatea unei efracții atunci când John și Mary sună? $P(Burglary|JohnCalls = true, MaryCalls = true)$



Inferență prin enumerare: exemplu

$P(Burglary|JohnCalls = true, MaryCalls = true)$

- Variabilele ascunse sunt *Earthquake* și *Alarm*.
  Conform ecuației (3),
  $\mathbf{P}(B|j,m) = \alpha \mathbf{P}(B,j,m) = \alpha \sum_e \sum_a \mathbf{P}(B,j,m,e,a)$
- Pentru $Burglary = true$, conform ecuației (2):
  $P(b|j,m) = \alpha \sum_e \sum_a P(b)P(e)P(a|b,e)P(j|a)P(m|a)$

  Termenul $P(b)$ const, $P(e)$ nu depinde de a:

$$
\begin{aligned}
P(b|j,m) &= \alpha P(b) \sum_e P(e) \sum_a P(a|b,e)P(j|a)P(m|a) \\
&= \alpha P(b) \sum_e P(e)[P(a|b,e)P(j|a)P(m|a) + \\
&\quad P(\neg a|b,e)P(j|\neg a)P(m|\neg a)] \\
&= \ldots \\
&= \alpha \times 0.00059224
\end{aligned}
$$

Inferență prin enumerare: exemplu

Pentru *Burglary* = *false*, conform ecuației (2):

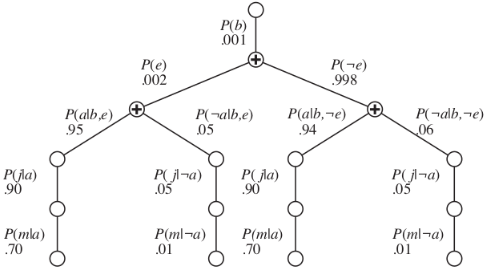
$$P(\neg b|j,m) = \alpha \sum_e \sum_a P(\neg b)P(e)P(a|\neg b,e)P(j|a)P(m|a)$$
$$= ...$$
$$= \alpha \times 0.0014919$$

$P(b|j,m) + P(\neg b|j,m) = 1 \rightarrow alpha = 479.8142$
$P(B|j,m) = \alpha\langle 0.00059224, 0.0014919\rangle \approx \langle 0.284, 0.716\rangle$.
Probabilitatea unei spargeri: 28%.

Obs: pentru a crește eficiența calculelor, se recomandă ca variabilele rămase să fie mai întâi sortate topologic, a.i. părinții să apară înaintea copiilor.
În acest caz, se vor putea descompune mai ușor sumele, scoțând în față factorii care nu depind de o anumită variabilă.

Exemplu

Procesul de evaluare:



Inferență prin enumerare

function ENUMERATION-ASK(X, $\mathbf{e}$, bn) **returns** a distribution over X
  inputs: X, the query variable
    $\mathbf{e}$, observed values for variables $\mathbf{E}$
    bn, a Bayes net with variables $\{X\} \cup \mathbf{E} \cup \mathbf{Y}$ /* $\mathbf{Y}$ = *hidden variables* */

  $\mathbf{Q}(X) \leftarrow$ a distribution over X, initially empty
  for each value x_i of X **do**
    $\mathbf{Q}(x_i) \leftarrow$ ENUMERATE-ALL(bn.VARS, $\mathbf{e}_{x_i}$)
      where $\mathbf{e}_{x_i}$ is $\mathbf{e}$ extended with $X = x_i$
  return NORMALIZE($\mathbf{Q}(X)$)

---

function ENUMERATE-ALL($vars$, $\mathbf{e}$) **returns** a real number
  if EMPTY?($vars$) **then return** 1.0
  $Y \leftarrow$ FIRST($vars$)
  if Y has value y in $\mathbf{e}$
    then return $P(y \mid parents(Y)) \times$ ENUMERATE-ALL(REST($vars$), $\mathbf{e}$)
    else return $\sum_y P(y \mid parents(Y)) \times$ ENUMERATE-ALL(REST($vars$), $\mathbf{e}_y$)
      where $\mathbf{e}_y$ is $\mathbf{e}$ extended with $Y = y$

Algoritmul evaluează arborii de expresii în manieră DFS. Complexitatea spațiu: $O(n)$, n variabile. Complexitatea timp: $O(2^n)$ pentru o rețea cu n variabile *bool*.

Algoritmul de eliminare a variabilelor (*variable elimination*)

- Obs: $P(j|a)P(m|a)$ și $P(j|\neg a)P(m|\neg a)$ sunt calculate de două ori, pentru fiecare valoare a lui e.

- Idee: realizează calculele de la dreapta către stânga (de jos în sus) și salvează rezultatele.

Variable elimination: exemplu

$$P(B|j,m) = \alpha \underbrace{P(b)}_{f_1(B)} \sum_e \underbrace{P(e)}_{f_2(E)} \sum_a \underbrace{P(a|B,e)}_{f_3(A,B,E)} \underbrace{P(j|a)}_{f_4(A)} \underbrace{P(m|a)}_{f_5(A)}$$

- Fiecare factor f_i este o matrice indexată de variabilele argument:
$f_4(A) = \begin{pmatrix} P(j|a) \\ P(j|\neg a) \end{pmatrix} = \begin{pmatrix} 0.90 \\ 0.05 \end{pmatrix}$, $f_5(A) = \begin{pmatrix} P(m|a) \\ P(m|\neg a) \end{pmatrix} = \begin{pmatrix} 0.70 \\ 0.01 \end{pmatrix}$.
$f_3(A,B,E)$ este o matrice $2x2x2$.

$$P(B|j,m) = \alpha f_1(B) \times \sum_e f_2(E) \times \sum_a f_3(A,B,E) \times f_4(A) \times f_5(A)$$

Variable elimination: exemplu

$$f_6(B,E) = \sum_a f_3(A,B,E) \times f_4(A) \times f_5(A)$$
$$= f_3(a,B,E) \times f_4(a) \times f_5(a) + f_3(\neg a,B,E) \times f_4(\neg a) \times f_5(\neg a)$$

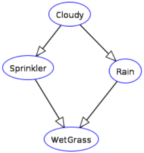
$$f_7(B) = \sum_e f_2(E) \times f_6(B,E)$$
$$= f_2(e) \times f_6(B,e) + f_2(\neg e) \times f_6(B,\neg e)$$

$$P(B|j,m) = \alpha f_1(B) \times f_7(B)$$

Inferență aproximativă

- Algoritmii exacți nu pot fi aplicați pentru rețelele complexe, cu sute de noduri. Inferența aproximativă poate crește viteza de calcul.

- Algoritmi de eșantionare (*sampling*) aleatoare pentru calculul probabilităților aposteriori
  Metode de eșantionare. Exemplu:



Considerăm ordonarea *Cloudy*, *Sprinkler*, *Rain*, *WetGrass*.
Eșantionăm din $\mathbf{P}(Cloudy) =< 0.5, 0.5 >$ valoarea *true*.
Eșantionăm din $\mathbf{P}(Sprinkler|Cloudy = true) =< 0.1, 0.9 >$ valoarea *false*, etc.

Pentru a calcula probabilitățile condiționale: *Rejection sampling*, *Likelihood weighting*.

Bibliografie

- S. Russell, P. Norvig. *Artificial Intelligence: A Modern Approach*. Ch. 13. Quantifying Uncertainty; Ch. 14. Probabilistic Reasoning
- Belief and Decision Networks `https://aispace.org/bayes/`