Rețele bayesiene

IA 2025/2026

Conținut

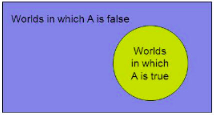
Introducere

- O rețea bayesiană modelează relații de dependență probabilistică între variabile
- Estimarea probabilității unui eveniment pe baza observațiilor
  Modelarea unor procese cauzale (vreme $\rightarrow$ trafic $\rightarrow$ întarzieri)
  Diagnostic medical (boala, simptome, teste)
  Predicția comportamentului utilizatorilor, etc.

Probabilități

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

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



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 $P(doubles) = 1/4.$

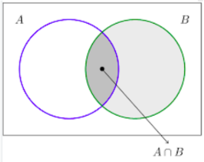
Probabilități necondiționate (*prior*)

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 **discrete**
  $P(Weather = sunny) = 0.6$
  $P(Weather = rain) = 0.1$
  $P(Weather = cloudy) = 0.29$
  $P(Weather = snow) = 0.01$

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

- Funcția densitate de probabilitate pentru a descrie distribuția unei var. aleatoare **continue**
  $P(NoonTemp = x) = Uniform_{[18C,26C]}(x)$ 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: probabilitățile tuturor combinațiilor de valori ale var.

  $\mathbf{P}(Weather, Cavity)$: o tabela 4 x 2
  Obs: în locul celor 8 ecuații, putem utiliza
  $\mathbf{P}(Weather, Cavity) = \mathbf{P}(Weather|Cavity)\mathbf{P}(Cavity)$.

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

Exemplu: Distribuția comună de probabilitate

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

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

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

Conținut

Distribuție marginală

Distribuția marginală peste o var. / o submulțime de 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: $\mathbf{P}(Cavity) = \sum_{\mathbf{z} \in \{Catch, Toothache\}} \mathbf{P}(Cavity, \mathbf{z})$

Prob. marginală $P(cavity) = 0.108 + 0.012 + 0.072 + 0.008 = 0.2$

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

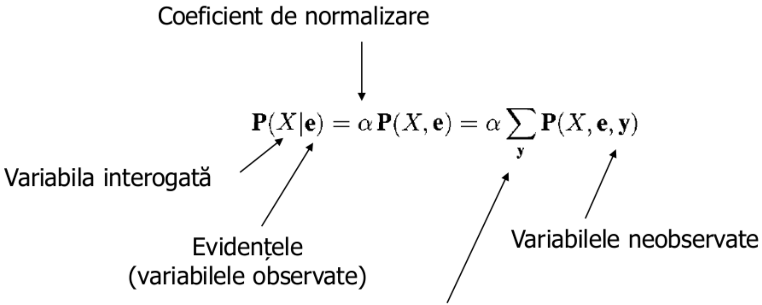
- Condiționare $\mathbf{P}(\mathbf{Y}) = \sum_{\mathbf{z}} \mathbf{P}(\mathbf{Y}|\mathbf{z}) P(\mathbf{z})$

Inferență

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

Procedura de inferență: fie variabila X, **E** lista de variabile evidență, **e** lista de valori observate, **Y** variabile neobservate

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



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

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

$$P(\neg cavity|toothache) = 0.4$$

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

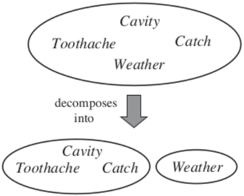
Obs: $1/P(toothache)$ const. de normalizare pentru distribuția $\mathbf{P}(Cavity|toothche)$.

Conținut

Independență

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

Exemplu: adăugăm variabila *Weather*.



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

Independență: exemplu

- $P(toothache, catch, cavity, cloudy) =$
  $P(cloudy|toothache, catch, cavity)P(toothache, catch, cavity)$ (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)
- $\mathbf{P}(I|E)$ probabilitatea **a posteriori** a ipotezei, dată fiind evidența (*posterior*)
- $\mathbf{P}(E|I)$ 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)

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

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:

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

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:

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

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.

Exemplul 2. Ion și Maria dau cu același ban

- 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

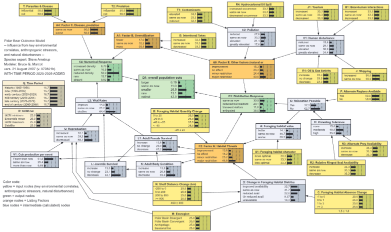
Conținut

Rețele bayesiene

- **Modele grafice probabilistice** care 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, modelarea mediului (populațiile de urși polari), etc.

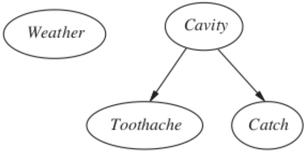


Rețea bayesiană

Este un digraf aciclic

- fiecare nod corespunde unei variabile aleatoare (eveniment)
- un arc de la X la Y (X este părintele lui Y): 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)

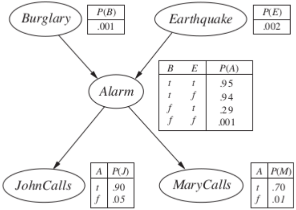
Topologia rețelei specifică relațiile de independență condiționată:



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

Rețele bayesiene: exemplu

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.

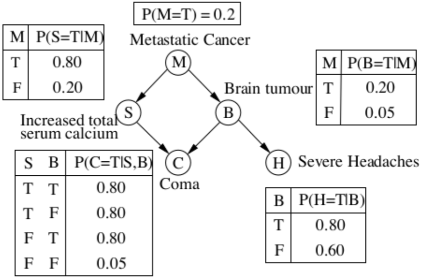


Efracțiile și cutremurele afectează probabilitatea declanșării alarmei. Alarma influențează probabilitatea ca John și Mary să sune.

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

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.



Obs: alăturat avem tabelele de probabilitate condiționată.

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:

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

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)**)

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

$\forall X_i$ 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*

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.

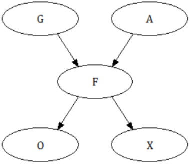
Sortare topologică: algoritmul lui Kahn

```
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), $S = \{O\}$, ...
   $\rightarrow L = \{G, A, F, O, X\}$

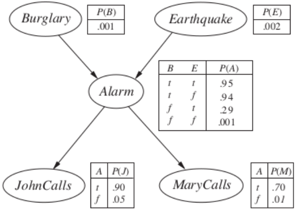
Rețele bayesiene: avantaje reprezentare

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.

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

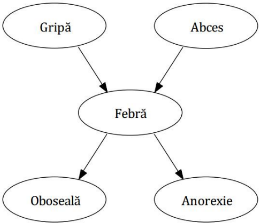
Burglary este independent de *JohnCalls* și *MaryCalls*, dat *Alarm* și *Earthquake*

Conținut

Inferența probabilităților marginale

- Calculează probabilitățile nodurilor, în **lipsa unor noduri evidență**
- Pentru un nod: calculăm suma **probabilităților condiționate**, înmulțite cu **probabilitățile părinților** (toate combinațiile posibile de valori ale părinților).

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

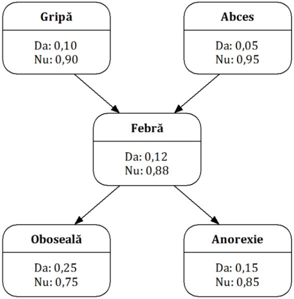
$$
\begin{aligned}
P(f) &= P(f|g,a)P(g)P(a) + P(f|g,\neg a)P(g)P(\neg a) \\
&\quad + P(f|\neg g,a)P(\neg g)P(a) + P(f|\neg g,\neg a)P(\neg g)P(\neg a) \\
&= 0.8 \cdot 0.1 \cdot 0.05 + 0.7 \cdot 0.1 \cdot 0.95 + 0.25 \cdot 0.9 \cdot 0.05 + 0.05 \cdot 0.9 \cdot 0.95 \\
&= 0.1245
\end{aligned}
$$

$$P(\neg f) = 1 - P(f) = 0.8755$$

Exemplu: Oboseală și Anorexie

$$P(o) = P(o|f)P(f) + P(o|\neg f)P(\neg f)$$
$$= 0.6 \cdot 0.1245 + 0.2 \cdot 0.8755 = 0.25$$
$$P(\neg o) = 1 - P(o) = 0.75$$

$$P(x) = P(x|f)P(f) + P(x|\neg f)P(\neg f)$$
$$= 0.5 \cdot 0.1245 + 0.1 \cdot 0.8755 = 0.15$$
$$P(\neg x) = 1 - P(x) = 0.85$$



Inferență în rețele bayesiene

- Sistem de **inferență probabilist**: calculăm distribuția de probabilitate a posteriori 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 a posteriori $\mathbf{P}(X|\mathbf{e}) =?$

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

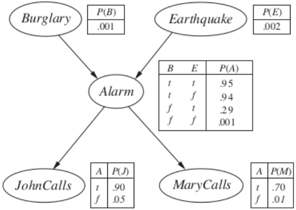
Inferență prin enumerare

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

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*.
  $\mathbf{P}(B|j,m) = \alpha \mathbf{P}(B,j,m) = \alpha \sum_e \sum_a \mathbf{P}(B,j,m,e,a)$

- 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):

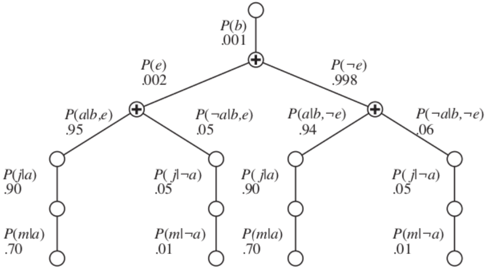
$$
\begin{aligned}
P(\neg b|j, m) &= \alpha \sum_e \sum_a P(\neg b)P(e)P(a|\neg b, e)P(j|a)P(m|a) \\
&= ... \\
&= \alpha \times 0.0014919
\end{aligned}
$$

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

Obs: pentru a eficientiza calculul, 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:



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 $2x2x2$.
  f_4 depinde doar de A deoarece $MaryCalls = true$.

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

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 aleatoare (**sampling**) pentru calculul **probabilităților condiționate**: *Rejection sampling*, *Likelihood weighting*.
  - eșantionăm dintr-o distribuție N eșantioane
  - aproximăm probabilitatea a posteriori $\hat{P}$ (converge la probabilitatea adevărată P)

Prior sampling

Pentru a genera un eveniment particular:

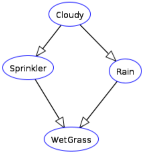
function PRIOR-SAMPLE(bn) **returns** an event sampled from the prior specified by bn
inputs: bn, a Bayesian network specifying joint distribution $\mathbf{P}(X_1, \ldots, X_n)$

$\mathbf{x} \leftarrow$ an event with n elements
foreach variable X_i **in** $X_1, \ldots, X_n$ **do**
$\mathbf{x}[i] \leftarrow$ a random sample from $\mathbf{P}(X_i \mid parents(X_i))$
return $\mathbf{x}$

Probabilitatea de a genera un eveniment particular:
$S_{PS}(x_1 \ldots x_n) = \prod_{i=1} P(x_i|parents(X_i)) = P(x_1 \ldots x_n)$

Prior sampling: 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*.
Eșantionăm din $\mathbf{P}(Rain|Cloudy = true) =< 0.8, 0.2 >$ valoarea *true*.
Eșantionăm din $\mathbf{P}(WetGrass|Sprinkler = false, Rain = true) =< 0.9, 0.1 >$ valoarea *true*.

$$S_{PS}(t, f, t, t) = 0.5 \times 0.9 \times 0.8 \times 0.9 = 0.324 = P(t, f, t, t)$$

Fie $N_{PS}(x_1 \ldots x_n)$ nr. de eșantioane generate pentru evenimentul $x_1, \ldots, x_n$.

$$\begin{aligned} lim_{N\to\infty}\hat{P}(x_1, \ldots, x_n) &= lim_{N\to\infty} N_{PS}(x_1, \ldots, x_n)/N \\ &= S_{PS}(x_1, \ldots, x_n) \\ &= P(x_1 \ldots x_n) \end{aligned}$$

Rejection sampling

- generăm eșantioane din distribuția apriori
  Dacă în 511 eșantioane din 1000 'Rain=true', atunci $\hat{P}(Rain = true) = 0.511$.
- respingem eșantioanele care nu se potrivesc cu evidența
- estimăm $\hat{P}(X = x|e)$

$$\hat{\mathbf{P}}(X|e) = \alpha N_{PS}(X, e) = \frac{N_{PS}(X, e)}{N_{PS}(e)}$$

Exemplu: estimăm $\mathbf{P}(Rain|Sprinkler = true)$
Din 100 eșantioane generate, 73 au 'Sprinkler = false' (respinse) și 27 au 'Sprinkler = true'; din cele 27, 8 au 'Rain = true' și 19 'Rain = false'.

$$\mathbf{P}(Rain|Sprinkler = true) \approx NORMALIZE(< 8, 19 >) =< 0.296, 0.704 >$$

Rejection sampling

```
function REJECTION-SAMPLING(X, e, bn, N) returns an estimate of P(X|e)
  inputs: X, the query variable
          e, observed values for variables E
          bn, a Bayesian network
          N, the total number of samples to be generated
  local variables: N, a vector of counts for each value of X, initially zero

  for j = 1 to N do
     x ← PRIOR-SAMPLE(bn)
     if x is consistent with e then
        N[x] ← N[x]+1 where x is the value of X in x
  return NORMALIZE(N)
```

Dezavantaj: multe eșantioane sunt respinse (nr. de eșantioane consistente scade exponențial cu nr. de var. evidență)

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/`