

PARTEA A II-A

În continuare ne propunem să calculăm probabilitatea $P(E_X^- | X)$, care intervine în formula (12), urmărind, în egală măsură, obținerea unei soluții recursive.

$$P(X | E) = \alpha P(E_X^- | X) \sum_{\mathbf{u}} P(X | \mathbf{u}) \prod_{i=1}^m P(\{U_i = u_i\} | E_{U_i \setminus X}) \quad (12)$$

Întrucât $E_{Y_i \setminus X}$ d-separă pe Y_i de toate celelalte dovezi din E_X^- , rezultă independența dovezilor din fiecare "casetă Y_i " față de celelalte, condiționat de X . Forma lui E_X^- este cea dată de formula (2). Datorită independenței dovezilor de sub X avem:

$$P(E_X^- | X) = \prod_{i=1}^n P(E_{Y_i \setminus X} | X) \quad (13)$$

Fie \mathbf{Y} vectorul fiilor Y_1, \dots, Y_n și fie $\mathbf{y} = (y_1, \dots, y_n)$ o realizare a acestuia. În cele ce urmează, vom ține cont de valorile fiilor lui $X, Y_i (i = \overline{1, n})$, dar va trebui să includem și părinții fiecărui Y_i . Fie \mathbf{Z}_i părinții lui Y_i , alții decât X și fie \mathbf{z}_i o atribuire de valori ale părinților.

Luând în considerație evenimentul sigur $\bigcup_{y_i} \{Y_i = y_i\}$ și respectiv

$$\bigcup_{\mathbf{z}_i} \{\mathbf{Z}_i = \mathbf{z}_i\}, \text{ avem:}$$

$$\begin{aligned}
P(E_{Y_i \setminus X} | X) &= P\left(E_{Y_i \setminus X} \cap \bigcup_{y_i} \{Y_i = y_i\} \cap \bigcup_{z_i} \{Z_i = z_i\} | X\right) = \\
&= \sum_{y_i} \sum_{z_i} P(E_{Y_i \setminus X}, \{Y_i = y_i\}, \{Z_i = z_i\} | X)
\end{aligned}$$

Aplicând, în formula anterioară, versiunea condiționată a formulei de înmulțire a probabilităților, obținem:

$$P(E_{Y_i \setminus X} | X) = \sum_{y_i} \sum_{z_i} P(\{Y_i = y_i, Z_i = z_i\} | X) P(E_{Y_i \setminus X} | X, \{Y_i = y_i, Z_i = z_i\}) \quad (14)$$

Mulțimea dovezilor $E_{Y_i \setminus X}$ reprezintă reuniunea a două submulțimi disjuncte și independente de variabile aleatoare:

$$E_{Y_i \setminus X} = E_{Y_i \setminus X}^+ \cup E_{Y_i}^-$$

Prin urmare, avem:

$$\begin{aligned}
P(E_{Y_i \setminus X} | X, \{Y_i = y_i, Z_i = z_i\}) &= P(E_{Y_i \setminus X}^+ | X, \{Y_i = y_i, Z_i = z_i\}) \cdot \\
&\cdot P(E_{Y_i}^- | X, \{Y_i = y_i, Z_i = z_i\}) \quad (15)
\end{aligned}$$

Dar $E_{Y_i}^-$ este o mulțime de variabile aleatoare independente de X și Z_i ,

iar $E_{Y_i \setminus X}^+$ constituie o mulțime de variabile aleatoare independente de X

și Y_i . În aceste condiții, repartițiile condiționate din (15) devin:

$$P(E_{Y_i \setminus X} | X, \{Y_i = y_i, Z_i = z_i\}) = P(E_{Y_i}^- | \{Y_i = y_i\}) \cdot P(E_{Y_i \setminus X}^+ | \{Z_i = z_i\}) \quad (16)$$

Cea de-a doua repartiție condiționată din membrul drept al lui (16) apare ca o repartiție a posteriori și poate fi exprimată cu formula lui Bayes (în versiune necondiționată):

$$P(E_{Y_i \setminus X}^+ | \{\mathbf{Z}_i = \mathbf{z}_i\}) = \frac{P(E_{Y_i \setminus X}^+) \cdot P(\{\mathbf{Z}_i = \mathbf{z}_i\} | E_{Y_i \setminus X}^+)}{P(\{\mathbf{Z}_i = \mathbf{z}_i\})} \quad (17)$$

Introducând (17) în (16) obținem:

$$\begin{aligned} P(E_{Y_i \setminus X} | X, \{Y_i = y_i, \mathbf{Z}_i = \mathbf{z}_i\}) &= \\ &= P(E_{Y_i}^- | \{Y_i = y_i\}) \cdot \frac{P(E_{Y_i \setminus X}^+) P(\{\mathbf{Z}_i = \mathbf{z}_i\} | E_{Y_i \setminus X}^+)}{P(\{\mathbf{Z}_i = \mathbf{z}_i\})} \end{aligned} \quad (18)$$

Introducând acum expresia lui $P(E_{Y_i \setminus X} | X, \{Y_i = y_i, \mathbf{Z}_i = \mathbf{z}_i\})$ dată de (18) în formula (14), rezultă:

$$\begin{aligned} P(E_{Y_i \setminus X} | X) &= \sum_{y_i} \sum_{\mathbf{z}_i} P(\{Y_i = y_i, \mathbf{Z}_i = \mathbf{z}_i\} | X) \cdot P(E_{Y_i}^- | \{Y_i = y_i\}) \cdot \\ &\quad \cdot \frac{P(E_{Y_i \setminus X}^+) P(\{\mathbf{Z}_i = \mathbf{z}_i\} | E_{Y_i \setminus X}^+)}{P(\{\mathbf{Z}_i = \mathbf{z}_i\})} \end{aligned} \quad (19)$$

Revenind acum la repartiția lui E_X^- condiționată de X , dată de formula (13) și luând în considerație (19), obținem:

$$P(E_X^- | X) = \prod_{i=1}^n \sum_{y_i} P(E_{Y_i}^- | \{Y_i = y_i\}) \sum_{\mathbf{z}_i} P(\{Y_i = y_i, \mathbf{Z}_i = \mathbf{z}_i\} | X) \cdot$$

$$\cdot \frac{P(E_{Y_i \setminus X}^+) P(\{\mathbf{Z}_i = \mathbf{z}_i\} | E_{Y_i \setminus X}^+)}{P(\{\mathbf{Z}_i = \mathbf{z}_i\})} \quad (20)$$

Aplicând în (20) versiunea condiționată a formulei de înmulțire a probabilităților, rezultă:

$$P(E_X^- | X) = \prod_{i=1}^n \sum_{y_i} P(E_{Y_i}^- | \{Y_i = y_i\}) \cdot \sum_{\mathbf{z}_i} \frac{P(E_{Y_i \setminus X}^+) P(\{\mathbf{Z}_i = \mathbf{z}_i\} | E_{Y_i \setminus X}^+)}{P(\{\mathbf{Z}_i = \mathbf{z}_i\})} \cdot P(\{\mathbf{Z}_i = \mathbf{z}_i\} | X) \cdot P(\{Y_i = y_i\} | X, \{\mathbf{Z}_i = \mathbf{z}_i\}) \quad (21)$$

Dar \mathbf{Z} și X sunt d-separate, ceea ce înseamnă că $P(\{\mathbf{Z}_i = \mathbf{z}_i\} | X) = P(\{\mathbf{Z}_i = \mathbf{z}_i\})$. Putem, de asemenea, înlocui pe $P(E_{Y_i \setminus X}^+)$ cu o constantă de normalizare β_i . În aceste condiții, egalitatea (21) devine:

$$P(E_X^- | X) = \prod_{i=1}^n \sum_{y_i} P(E_{Y_i}^- | \{Y_i = y_i\}) \sum_{\mathbf{z}_i} \beta_i P(\{\mathbf{Z}_i = \mathbf{z}_i\} | E_{Y_i \setminus X}^+) \cdot P(\{Y_i = y_i\} | X, \{\mathbf{Z}_i = \mathbf{z}_i\})$$

Combinând toți β_i într-o unică constantă de normalizare β și ținând cont de semnificația lui $E_{Y_i \setminus X}^+$, obținem:

$$P(E_X^- | X) = \beta \prod_i \sum_{y_i} P(E_{Y_i}^- | y_i) \sum_{\mathbf{z}_i} P(y_i | X, \mathbf{z}_i) \prod_j P(z_{ij} | E_{Z_{ij} \setminus Y_i}) \quad (22)$$

Se observă că fiecare dintre termenii expresiei finale date de (22) este ușor de evaluat:

- $P(E_{Y_i}^- | y_i)$ este o secvență recursivă a calculării lui $P(E_X^- | X)$;
- $P(y_i | X, \mathbf{z}_i)$ se citesc din tabelul de probabilități condiționate asociat lui Y_i ;
- $P(z_{ij} | E_{Z_{ij} \setminus Y_i})$ reprezintă o secvență recursivă a problemei inițiale, adică a calculării lui $P(X | E)$, mai precis a lui $P(X | E_{X \setminus V})$.

Problema transformării acestor calcule într-un *algorithm* este relativ simplă. Vor fi necesare două subrutine de bază, care vor calcula $P(X | E_{X \setminus V})$ și respectiv $P(E_{X \setminus V}^- | X)$, unde prin V am notat variabila care desemnează dovezile exceptate atunci când sunt luate în considerație cele conectate la nodul X .

În Algoritmul 5.1 subrutina SUPORT-EXCEPT(X, V) calculează $P(X | E_{X \setminus V})$ folosind o generalizare a ecuației (12), în timp ce subrutina DOVEZI-EXCEPT(X, V) calculează $P(E_{X \setminus V}^- | X)$ folosind o generalizare a ecuației (22).

Pentru a simplifica prezentarea, se va presupune că rețeaua este fixată și deja "aprovizionată" cu dovezi, precum și că variabilele dovezi satisfac predicatul DOVEZI?. Probabilitățile $P(X | \mathbf{U})$, unde \mathbf{U} reprezintă vectorul părinților lui X , sunt disponibile în tabelul de probabilități condiționate asociat lui X . Calculul expresiilor $\alpha \dots$ și $\beta \dots$ se face prin normalizare.

În general, un agent primește valori ale variabilelor dovezi prin intermediul senzorilor săi sau în urma unor raționamente și se interesează (întreabă) despre posibilele valori ale altor variabile, astfel încât să poată decide ce acțiune trebuie întreprinsă în continuare. Algoritmul 5.1 calculează distribuția de probabilitate condiționată pentru o variabilă de interogare dată. El reprezintă un algoritm de tip înlănțuire-înapoi pentru rezolvarea interogărilor probabiliste asupra unui polyarbore.

Algoritmul 5.1

```

function INTEROGARE-REȚEA  $(X)$  return
    o distribuție de probabilitate a valorilor lui  $X$ 
input:  $X$ , o variabilă aleatoare
    SUPORT-EXCEPT  $(X, \text{nul})$ 

function SUPORT-EXCEPT  $(X, V)$  return  $P(X | E_{X \setminus V})$ 

    if DOVEZI ?  $(X)$  then return distribuția observată pentru  $X$ 
    else
        calculează  $P(E_{X \setminus V}^- | X) = \text{DOVEZI-EXCEPT}(X, V)$ 
         $U \leftarrow \text{PĂRINȚI}[X]$ 
        if  $U$  este vid
            then return  $\propto P(E_{X \setminus V}^- | X) P(X)$ 
        else
            pentru fiecare  $U_i$  in  $U$  do

```

calculează și memorează $P(U_i | E_{U_i \setminus X}) =$
 $= \text{SUPPORT-EXCEPT}(U_i, X)$
 $\alpha P(E_{X \setminus V}^- | X) \sum_{\mathbf{u}} P(X | \mathbf{u}) \prod_i P(U_i | E_{U_i \setminus X})$
return

function DOVEZI-EXCEPT (X, V) **return** $P(E_{X \setminus V}^- | X)$
 $\mathbf{Y} \leftarrow \text{FII}[X] - V$
if \mathbf{Y} este vid
 then return o repartiție uniformă
else
 pentru fiecare Y_i **in** \mathbf{Y} **do**
 calculează $P(E_{Y_i}^- | y_i) = \text{DOVEZI-EXCEPT}$
 (Y_i, nul)
 $\mathbf{Z}_i \leftarrow \text{PĂRINȚI}[Y_i] - X$
 pentru fiecare Z_{ij} **in** \mathbf{Z}_i
 calculează $P(Z_{ij} | E_{Z_{ij} \setminus Y_i}) =$
 $= \text{SUPPORT-EXCEPT}(Z_{ij}, Y_i)$
 return

$\beta \prod_i \sum_{y_i} P(E_{Y_i}^- | y_i) \sum_{\mathbf{z}_i} P(y_i | X, \mathbf{z}_i) \prod_j P(z_{ij} | E_{Z_{ij} \setminus Y_i})$

□

Calcululele efectuate de Algoritmul 5.1 presupun apeluri recursive care se extind de la X de-a lungul tuturor drumurilor din rețea. Recursivitatea se încheie atunci când sunt atinse noduri-dovezi, noduri-rădăcină (care nu au părinți) și noduri-frunză (care nu au fii). Fiecare apel recursiv exclude nodul de la care s-a produs apelul, astfel încât fiecare nod din arbore este tratat o singură dată. Prin urmare, algoritmul este liniar raportat la numărul de noduri ale rețelei. Algoritmul se comportă în acest fel deoarece rețeaua dată reprezintă un *polyarbore*. Dacă ar exista mai mult de un singur drum între o pereche de noduri, atunci fie procedurile recursive ar lua în calcul aceleași dovezi de mai multe ori, fie execuția nu s-ar încheia într-un număr finit de pași.

Algoritmul prezentat, de tip înlănțuire-înapoi, este cel mai simplu algoritm pentru polyarbori. Principalul neajuns al unui asemenea algoritm este acela că el calculează repartiția de probabilitate pentru o unică variabilă. Dacă se dorește determinarea repartițiilor a posteriori pentru toate variabilele care nu constituie variabile dovezi, atunci programul implementând Algoritmul 5.1 ar trebui executat pentru fiecare dintre acestea, ceea ce ar mări timpul de execuție. În acest caz este preferabilă o abordare de tip înlănțuire-înainte, care pleacă de la variabilele dovezi. În cazul unei contabilizări adecvate, calculele pot fi efectuate într-un timp liniar. Versiunea algoritmului care folosește înlănțuirea-înainte poate fi privită ca reprezentând o "propagare a mesajelor" prin rețea. Din această cauză implementarea pe calculatoare paralele este relativ simplă și pot fi făcute analogii interesante cu propagarea mesajelor între neuronii din creier.