PARTEA A II-A

În continuare ne propunem să calculăm probabilitatea $P\left(E_X^- \mid X\right)$, care intervine în formula (12), urmărind, în egală măsură, obţinerea unei soluţii recursive.

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

Întrucât $E_{Y_i \setminus X}$ d-separă pe Y_i de toate celelalte dovezi din E_X^-, rezultă independenţa dovezilor din fiecare ”casetă Y_i” faţă de celelalte, condiţionat de X. Forma lui E_X^- este cea dată de formula (2). Datorită independenţei dovezilor de sub X avem:

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

Fie $\mathbf{Y}$ vectorul fiilor $Y_1, \ldots, Y_n$ şi fie $\mathbf{y} = \left(y_1, \ldots, y_n\right)$ o realizare a acestuia. În cele ce urmează, vom ţine cont de valorile fiilor lui $X, Y_i \left(i = \overline{1,n}\right)$, dar va trebui să includem şi părinţii fiecărui Y_i. Fie $\mathbf{Z}_i$ părinţii lui Y_i, alţii decât X şi fie $\mathbf{z}_i$ o atribuire de valori ale părinţilor. Luând în consideraţie evenimentul sigur $\bigcup_{\mathbf{y}_i} \left\{Y_i = y_i\right\}$ şi respectiv $\bigcup_{\mathbf{z}_i} \left\{\mathbf{Z}_i = \mathbf{z}_i\right\}$, avem:

$$P\left(E_{Y_i \setminus X} \mid X\right) = P\left(E_{Y_i \setminus X} \bigcap \bigcup_{y_i} \left\{Y_i = y_i\right\} \bigcap \bigcup_{\mathbf{z}_i} \left\{\mathbf{Z}_i = \mathbf{z}_i\right\} \mid X\right) =$$
$$= \sum_{y_i} \sum_{\mathbf{z}_i} P\left(E_{Y_i \setminus X}, \left\{Y_i = y_i\right\}, \left\{\mathbf{Z}_i = \mathbf{z}_i\right\} \mid X\right)$$

Aplicând, în formula anterioară, versiunea condiţionată a formulei de înmulţire a probabilităţilor, obţinem:

$$P\left(E_{Y_i \setminus X} \mid X\right) = \sum_{y_i} \sum_{\mathbf{z}_i} P\left(\{Y_i = y_i, \mathbf{Z}_i = \mathbf{z}_i\} \mid X\right) P\left(E_{Y_i \setminus X} \mid X, \{Y_i = y_i, \mathbf{Z}_i = \mathbf{z}_i\}\right) \quad (2)$$

Mulţimea dovezilor $E_{Y_i \setminus X}$ reprezintă reuniunea a două submulţimi disjuncte şi independente de variabile aleatoare:

$$E_{Y_i \setminus X} = E^{+}_{Y_i \setminus X} \bigcup E^{-}_{Y_i}$$

Prin urmare, avem:

$$P\left(E_{Y_i \setminus X} \mid X, \{Y_i = y_i, \mathbf{Z}_i = \mathbf{z}_i\}\right) = P\left(E^{+}_{Y_i \setminus X} \mid X, \{Y_i = y_i, \mathbf{Z}_i = \mathbf{z}_i\}\right) \cdot$$
$$\cdot P\left(E^{-}_{Y_i} \mid X, \{Y_i = y_i, \mathbf{Z}_i = \mathbf{z}_i\}\right) \quad (3)$$

Dar $E^{-}_{Y_i}$ este o mulţime de variabile aleatoare independente de X şi $\mathbf{Z}_i$, iar $E^{+}_{Y_i \setminus X}$ constituie o mulţime de variabile aleatoare independente de X şi Y_i. În aceste condiţii, repartiţiile condiţionate din (15) devin:

$$P\left(E_{Y_i \setminus X} \mid X, \{Y_i = y_i, \mathbf{Z}_i = \mathbf{z}_i\}\right) = P\left(E^{-}_{Y_i} \mid \{Y_i = y_i\}\right) \cdot P\left(E^{+}_{Y_i \setminus X} \mid \{\mathbf{Z}_i = \mathbf{z}_i\}\right) \quad (4)$$

Cea de-a doua repartiţie condiţionată din membrul drept al lui (16) apare ca o repartiţie a posteriori şi poate fi exprimată cu formula lui Bayes (în versiune necondiţionată):

$$P\left(E^{+}_{Y_i \setminus X} \mid \{\mathbf{Z}_i = \mathbf{z}_i\}\right) = \frac{P\left(E^{+}_{Y_i \setminus X}\right) \cdot P\left(\{\mathbf{Z}_i = \mathbf{z}_i\} \mid E^{+}_{Y_i \setminus X}\right)}{P\left(\{\mathbf{Z}_i = \mathbf{z}_i\}\right)} \quad (5)$$

Introducând (17) în (16) obţinem:

$$P\left(E_{Y_i \setminus X} \mid X, \{Y_i = y_i, \mathbf{Z}_i = \mathbf{z}_i\}\right) =$$

$$= P\left(E_{Y_i}^- \mid \{Y_i = y_i\}\right) \cdot \frac{P\left(E_{Y_i\setminus X}^+\right) P\left(\{\mathbf{Z}_i = \mathbf{z}_i\} \mid E_{Y_i\setminus X}^+\right)}{P\left(\{\mathbf{Z}_i = \mathbf{z}_i\}\right)} \qquad (6)$$

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

$$P\left(E_{Y_i\setminus X} \mid X\right) = \sum_{y_i} \sum_{\mathbf{z}_i} P\left(\{Y_i = y_i, \mathbf{Z}_i = \mathbf{z}_i\} \mid X\right) \cdot P\left(E_{Y_i}^- \mid \{Y_i = y_i\}\right) \cdot$$

$$\cdot \frac{P\left(E_{Y_i\setminus X}^+\right) P\left(\{\mathbf{Z}_i = \mathbf{z}_i\} \mid E_{Y_i\setminus X}^+\right)}{P\left(\{\mathbf{Z}_i = \mathbf{z}_i\}\right)} \qquad (7)$$

Revenind acum la repartiţia lui E_X^- condiţionată de X, dată de formula (13) şi luând în consideraţie (19), obţinem:

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

$$\cdot \frac{P\left(E_{Y_i\setminus X}^+\right) P\left(\{\mathbf{Z}_i = \mathbf{z}_i\} \mid E_{Y_i\setminus X}^+\right)}{P\left(\{\mathbf{Z}_i = \mathbf{z}_i\}\right)} \qquad (8)$$

Aplicând în (20) versiunea condiţionată a formulei de înmulţire a probabilităţilor, rezultă:

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

$$\cdot P\left(\{\mathbf{Z}_i = \mathbf{z}_i\} \mid X\right) \cdot P\left(\{Y_i = y_i\} \mid X, \{\mathbf{Z}_i = \mathbf{z}_i\}\right) \qquad (9)$$

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

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

Combinând toţi β_i într-o unică constantă de normalizare β şi ţinând cont de semnificaţia lui $E_{Y_i\setminus X}^+$, obţinem:

$$P\left(E_X^- \mid X\right)=\beta\prod_{i}\sum_{y_i}P\left(E_{Y_i}^- \mid y_i\right)\sum_{\mathbf{z}_i}P\left(y_i \mid X,\mathbf{z}_i\right)\prod_{j}P\left(z_{ij} \mid E_{Z_{ij}\setminus Y_i}\right) \qquad (10)$$

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

- $P\left(E_{Y_i}^- \mid y_i\right)$ este o secvenţă recursivă a calculării lui $P\left(E_X^- \mid X\right)$;
- $P\left(y_i \mid X,\mathbf{z}_i\right)$ se citesc din tabelul de probabilităţi condiţionate asociat lui Y_i;
- $P\left(z_{ij} \mid E_{Z_{ij}\setminus Y_i}\right)$ reprezintă o secvenţă recursivă a problemei iniţiale, adică a calculării lui $P\left(X \mid E\right)$, mai precis a lui $P\left(X \mid E_{X\setminus}\right)$.

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

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

Pentru a simplifica prezentarea, se va presupune că reţeaua este fixată şi deja ”aprovizionată” cu dovezi, precum şi că variabilele dovezi satisfac predicatul DOVEZI?. Probabilităţile $P(X \mid \mathbf{U})$, unde $\mathbf{U}$ reprezintă vectorul părinţilor lui X, sunt disponibile în tabelul de probabilităţi condiţionate asociat lui X. Calculul expresiilor $\alpha\ldots$ şi $\beta\ldots$ se face prin normalizare.

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

<u>Algoritmul 5.1</u>

function INTEROGARE-REŢEA(X) **return**

o distribuţie de probabilitate a valorilor lui X

input: X, o variabilă aleatoare

SUPORT-EXCEPT(X, nul)

function SUPORT-EXCEPT(X,V) **return** $P(X \mid E_{X\backslash V})$

if DOVEZI ? (X) **then return** distribuţia observată pentru X

else

calculează $P(E^{-}_{X\backslash V} \mid X)$ = DOVEZI-EXCEPT (X,V)

$\mathbf{U} \leftarrow$ PĂRINŢI $[X]$

if $\mathbf{U}$ este vid

then return $\alpha\, P\left(E^-_{X\setminus V} \mid X\right) P(X)$

else

pentru fiecare U_i **in** U **do**

calculează şi memorează $P\left(U_i \mid E_{U_i \setminus X}\right) =$

$=$SUPORT-EXCEPT $\left(U_i, X\right)$

return $\alpha\, P\left(E^-_{X\setminus V} \mid X\right) \sum_{\mathbf{u}} P(X \mid \mathbf{u}) \prod_i P\left(U_i \mid E_{U_i \setminus X}\right)$

function DOVEZI-EXCEPT (X, V) **return** $P\left(E^-_{X\setminus V} \mid X\right)$

$\mathbf{Y} \leftarrow$FII $[X] - V$

if $\mathbf{Y}$ este vid

then return o repartiţie uniformă

else

pentru fiecare Y_i **in** $\mathbf{Y}$ **do**

calculează $P\left(E^-_{Y_i} \mid y_i\right) =$DOVEZI-EXCEPT$\left(Y_i,\ \text{nul}\right)$

$\mathbf{Z}_i \leftarrow$PĂRINŢI $[Y_i] - X$

pentru fiecare Z_{ij} **in** $\mathbf{Z}_i$

calculează $P\left(Z_{ij} \mid E_{Z_{ij} \setminus Y_i}\right) =$

$=$SUPORT-EXCEPT$\left(Z_{ij}, Y_i\right)$

return $\beta \prod_i \sum_{y_i} P\left(E^-_{Y_i} \mid y_i\right) \sum_{\mathbf{z}_i} P\left(y_i \mid X, \mathbf{z}_i\right) \prod_j P\left(z_{ij} \mid E_{Z_{ij} \setminus Y_i}\right)$

❑

Calculele efectuate de Algoritmul 5.1 presupun apeluri recursive care se extind de la X de-a lungul tuturor drumurilor din reţea. Recursivitatea se încheie atunci când sunt atinse noduri-dovezi, noduri-rădăcină (care nu au părinţi) şi 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 reţelei. Algoritmul se comportă în acest fel deoarece reţeaua 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 aceleaşi dovezi de mai multe ori, fie execuţia nu s-ar încheia într-un număr finit de paşi.

Algoritmul prezentat, de tip înlănţuire-înapoi, este cel mai simplu algoritm pentru polyarbori. Principalul neajuns al unui asemenea algoritm este acela că el calculează repartiţia de probabilitate pentru o unică variabilă. Dacă se doreşte determinarea repartiţiilor 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 execuţie. În acest caz este preferabilă o abordare de tip înlănţuire-înainte, care pleacă de la variabilele dovezi. În cazul unei contabilizări adecvate, calculele pot fi efectuate într-un timp liniar. Versiunea algoritmului care foloseşte înlănţuirea-înainte poate fi privită ca reprezentând o ”propagare a mesajelor” prin reţea. Din această cauză implementarea pe calculatoare paralele este relativ simplă şi pot fi făcute analogii interesante cu propagarea mesajelor între neuronii din creier.