### Un algoritm pentru răspunsul la interogări

În cele ce urmează, ne propunem să stabilim un algoritm de calcul al probabilităților condiționate a posteriori ale variabilelor de interogare. Algoritmul găsit va fi de tip înlănțuire înapoi prin faptul că pleacă de la variabila de interogare şi urmează drumurile de la acel nod până la nodurile dovezi. Datorită complicațiilor care se pot naşte atunci când două drumuri diferite converg la acelaşi nod, algoritmul pe care îl vom prezenta se referă numai la *rețele unic conectate*, cunoscute şi sub denumirea de **polyarbori**. Reamintim faptul că, în astfel de rețele, există cel mult un drum nedirecționat între oricare două noduri ale rețelei. Algoritmii pentru rețele generale, asupra cărora nu ne vom opri în cadrul restrâns al acestui curs, vor folosi algoritmii referitori la polyarbori ca principală subrutină.

Fig. 5.4 prezintă o rețea generică unic conectată. În această rețea nodul $X$ are părinții $\mathbf{U} = U_1 \ldots U_m$ şi fiii $\mathbf{Y} = Y_1 \ldots Y_n$. Corespunzător fiecărui fiu şi fiecărui părinte a fost desenat un dreptunghi care include toți descendenții nodului şi toți strămoşii lui (cu excepția lui $X$). Proprietatea de *unică conectare* înseamnă că toate dreptunghiurile sunt disjuncte şi că nu există legături care să le conecteze între ele. Se presupune că $X$ este variabila de interogare şi că există o mulțime $E$ de variabile dovezi[1]. Se urmăreşte calcularea probabilității condiționate

$$P(X \mid E)$$

[1] $E = \{E_X^-, E_X^+\} = \{U_1, \ldots, U_m, Y_1, \ldots, Y_n\}$, mulțime de variabile aleatoare discrete.

Aceasta rețea este partiționată în conformitate cu părinții şi cu fiii variabilei de interogare $X$. Pentru a concepe un algoritm, va fi util să ne putem referi la diferite porțiuni ale dovezilor, astfel:

- $E_X^+$ reprezintă **suportul cauzal** pentru $X$ - variabilele dovezi aflate ”deasupra” lui $X$, care sunt conectate la $X$ prin intermediul părinților săi.
- $E_X^-$ reprezintă **suportul probatoriu** pentru $X$ - variabilele dovezi aflate ”dedesubtul” lui $X$ şi care sunt conectate la $X$ prin intermediul fiilor săi.

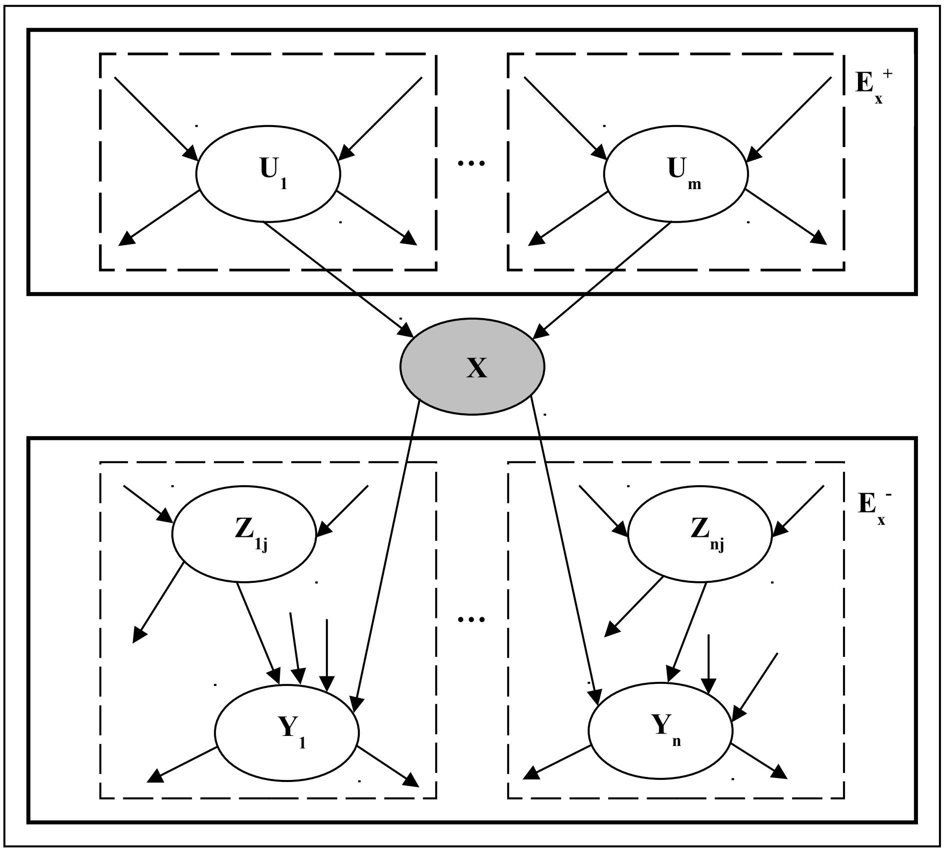


**Fig. 5.4**

Notația $E_{U_i \backslash X}$ va fi folosită pentru a se face referire la toate dovezile conectate cu nodul $U_i$, mai puțin cele prin drumul de la $X$. Cu această ultimă notație, putem partiționa pe $E_X^+$ în $E_{U_1 \backslash X}, \ldots, E_{U_m \backslash X}$ şi deci putem exprima pe $E_X^+$ sub forma următoare:

$$E_X^+ = \bigcup_{i=1}^{m} E_{U_i \backslash X}$$

În mod similar, $E_{Y_i \backslash X}$ semnifică mulțimea tuturor dovezilor conectate la $Y_i$ prin intermediul părinților săi, *cu excepția* lui $X$. Mulțimea $E_X^-$ poate fi deci partiționată în mulțimile $E_{Y_1 \backslash X}, \ldots, E_{Y_n \backslash X}$ şi este de forma

$$E_X^- = \bigcup_{i=1}^{n} E_{Y_i \backslash X}$$

Se observă că mulțimea tuturor dovezilor $E$ poate fi scrisă ca $E_X$ (toate dovezile conectate la $X$) şi ca $E_{X \backslash}$ (toate dovezile conectate la $X$, fără excepție).

Strategia generală pentru calculul lui $P(X \mid E)$ este atunci următoarea:

- Exprimă $P(X \mid E)$ în termenii contribuțiilor lui $E_X^+$ şi $E_X^-$.
- Calculează contribuția mulțimii $E_X^+$ calculând efectul ei asupra părinților lui $X$ şi apoi transmițând acest efect lui $X$. (Calcularea efectului asupra fiecărui părinte al lui $X$ este o secvență recursivă a problemei calculării efectului asupra lui $X$).

- Calculează contribuția mulțimii $E_X^-$ calculând efectul ei asupra fiilor lui $X$ şi apoi transmițând acest efect lui $X$. (Calcularea efectului asupra fiecărui fiu al lui $X$ reprezintă o secvență recursivă a problemei calculării efectului asupra lui $X$).

Totalitatea dovezilor, $E$, constă din dovezile aflate ”deasupra” lui $X$ şi din cele aflate ”dedesubtul” lui $X$, întrucât s-a făcut presupunerea că $X$ însuşi nu se află în $E$. Prin urmare, avem

$$P(X \mid E) = P(X \mid E_X^-, E_X^+)$$

Pentru a separa contribuțiile lui $E_X^+$ şi $E_X^-$, vom aplica versiunea condiționată a regulii lui Bayes, păstrând pe $E_X^-$ ca dovadă fixată în fundal:

$$P(X \mid E_X^-, E_X^+) = \frac{P(X \mid E_X^+)\, P(E_X^- \mid X, E_X^+)}{P(E_X^- \mid E_X^+)}$$

Întrucât $X$ d-separă în cadrul rețelei pe $E_X^+$ de $E_X^-$, putem folosi independența condiționată pentru a simplifica al doilea termen al numărătorului. De asemenea, putem trata $1/P(E_X^- \mid E_X^+)$ ca pe o constantă de normalizare, obținând:

$$P(X \mid E) = \alpha\, P(X \mid E_X^+)\, P(E_X^- \mid X)$$

Prin urmare, este necesar să calculăm cei doi termeni $P(X \mid E_X^+)$ şi $P(E_X^- \mid X)$. Vom începe prin tratarea primului, care se calculează relativ uşor.

Vom calcula $P\left(X \mid E_X^+\right)$ luând în considerație toate configurațiile posibile ale părinților lui $X$, precum şi cât de probabile sunt acestea fiind dată mulțimea $E_X^+$. În cazul fiecărei configurații date, probabilitatea lui $X$ se cunoaşte direct din tabelul probabilităților condiționate.

Fie $\mathbf{U}$ vectorul părinților $U_1,\ldots,U_m$ şi fie $\mathbf{u}$ o atribuire de valori pentru aceştia[2]. În calculele care urmează vom folosi faptul că $E_{U_i \setminus X}$ d-separă pe $U_i$ de toate celelalte dovezi din $E_X^+$. Ținând cont de acest fapt şi de formula (1), obținem egalitatea:

$$P\left(\left\{U_i = u_i\right\} \mid E_X^+\right) = P\left(\left\{U_i = u_i\right\} \mid E_{U_i \setminus X}\right) \quad i = \overline{1,m}$$

Luând în considerație evenimentul sigur, putem exprima probabilitatea $P\left(X \mid E_X^+\right)$ sub următoarea formă:

$$P\left(X \mid E_X^+\right) = P\left(X \cap \bigcup_{(u_1,\ldots,u_m)} \left\{U_1 = u_1,\ldots,U_m = u_m\right\} \mid E_X^+\right)$$

Întrucât membrul drept din (5) reprezintă probabilitatea unei reuniuni de mulțimi disjuncte, avem:

$$P\left(X \mid E_X^+\right) = \sum_{(u_1,\ldots,u_m)} P\left(X \cap \left\{U_1 = u_1,\ldots,U_m = u_m\right\} \mid E_X^+\right)$$

Aplicând, în continuare, versiunea condiționată a formulei de înmulțire a probabilităților membrului drept din (6), obținem:

$$P\left(X \mid E_X^+\right) = \sum_{(u_1,\ldots,u_m)} P\left(\left\{U_1 = u_1,\ldots,U_m = u_m\right\} \mid E_X^+\right) \cdot$$

---

[2] Spre exemplu, dacă există doi părinți Booleeni, $U_1$ şi $U_2$, atunci $\mathbf{u}$ trece peste patru atribuiri posibile, dintre care una este [true, false].

$$\cdot P\left(X \mid \left\{U_1 = u_1, \ldots, U_m = u_m\right\}, E_X^+\right)$$

Probabilitatea $P\left(\left\{U_1 = u_1, \ldots, U_m = u_m\right\} \mid E_X^+\right)$, care intervine în membrul drept al formulei (7), poate fi explicitată dacă se ține cont de independența variabilelor aleatoare $U_1, \ldots, U_m$, precum şi de faptul că $E_{U_i \setminus X}$ d-separă pe $U_i$ de toate celelalte dovezi din $E_X^+$, care a fost partiționat în $E_{U_1 \setminus X}, \ldots, E_{U_m \setminus X}$. Întrucât $U_1, \ldots, U_m$ sunt independente,

$$P\left(\left\{U_1 = u_1, \ldots, U_m = u_m\right\} \mid E_X^+\right) = \prod_{i=1}^{m} P\left(\left\{U_i = u_i\right\} \mid E_X^+\right)$$

şi, ținând cont de (1), avem:

$$\prod_{i=1}^{m} P\left(\left\{U_i = u_i\right\} \mid E_X^+\right) = \prod_{i=1}^{m} P\left(\left\{U_i = u_i\right\} \mid E_{U_i \setminus X}\right)$$

Cea de-a doua probabilitate condiționată care intervine în membrul drept al formulei (7) poate fi explicitată ținându-se cont de faptul că $U$ d-separă pe $X$ de $E_X^+$:

$$P\left(X \mid \left\{U_1 = u_1, \ldots, U_m = u_m\right\}, E_X^+\right) = P\left(X \mid \left\{U_1 = u_1, \ldots, U_m = u_m\right\}\right)$$

Ținând cont de (7), (8), (9) şi (10), obținem:

$$P\left(X \mid E_X^+\right) = \sum_{(u_1, \ldots, u_m)} P\left(X \mid \left\{U_1 = u_1, \ldots, U_m = u_m\right\}\right) \prod_{i=1}^{m} P\left(\left\{U_i = u_i\right\} \mid E_{U_i \setminus X}\right)$$

Introducând expresia lui $P\left(X \mid E_X^+\right)$ dată de (11) în formula (3), rezultă:

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

Ecuația (12) sugerează deja un algoritm. Astfel, $P(X \mid \mathbf{u})$ este repartiția lui $X$ condiționată de realizarea $(U_1 = u_1, \ldots, U_m = u_m)$. Valoarea acestei probabilități poate fi luată din tabelul de probabilități condiționate asociat lui $X$. Calculul fiecărei probabilități $P(\{U_i = u_i\} \mid E_{U_i \setminus X})$ reprezintă o secvență recursivă a problemei inițiale, aceea de a calcula $P(X \mid E)$, adică $P(X \mid E_{X\setminus})$. Vom mai nota aici faptul că mulțimea variabilelor dovezi care intervin în apelarea recursivă reprezintă o submulțime a celor din apelarea inițială, ceea ce constituie un indiciu că procedura de calcul se va termina într-un număr finit de pași, ea reprezentând într-adevăr un algoritm.

În continuare ne propunem să calculăm probabilitatea $P(E_X^- \mid X)$, care intervine în formula (12), urmărind, în egală măsură, obținerea unei soluții recursive. – **CURSUL URMATOR, la sfarsitul caruia vom putea formula un algoritm pentru raspunsul la interogari.**