Un algoritm pentru răspunsul la interogări

În cele ce urmează, ne propunem să stabilim un algoritm de calcul al probabilităţilor condiţionate a posteriori ale variabilelor de interogare. Algoritmul găsit va fi de tip înlănţuire înapoi prin faptul că pleacă de la variabila de interogare şi urmează drumurile de la acel nod până la nodurile dovezi. Datorită complicaţiilor care se pot naşte atunci când două drumuri diferite converg la acelaşi nod, algoritmul pe care îl vom prezenta se referă numai la *reţele unic conectate*, cunoscute şi sub denumirea de **polyarbori**. Reamintim faptul că, în astfel de reţele, există cel mult un drum nedirecţionat între oricare două noduri ale reţelei. Algoritmii pentru reţele 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 reţea generică unic conectată. În această reţea nodul X are părinţii $\mathbf{U} = U_1 \ldots U_m$ şi fiii $\mathbf{Y} = Y_1 \ldots Y_n$. Corespunzător fiecărui fiu şi fiecărui părinte a fost desenat un dreptunghi care include toţi descendenţii nodului şi toţi strămoşii lui (cu excepţia 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 mulţime E de variabile dovezi[1]. Se urmăreşte calcularea probabilităţii condiţionate

$$P(X \mid E)$$

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

Aceasta reţea este partiţionată în conformitate cu părinţii şi cu fiii variabilei de interogare X. Pentru a concepe un algoritm, va fi util să ne putem referi la diferite porţiuni ale dovezilor, astfel:

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

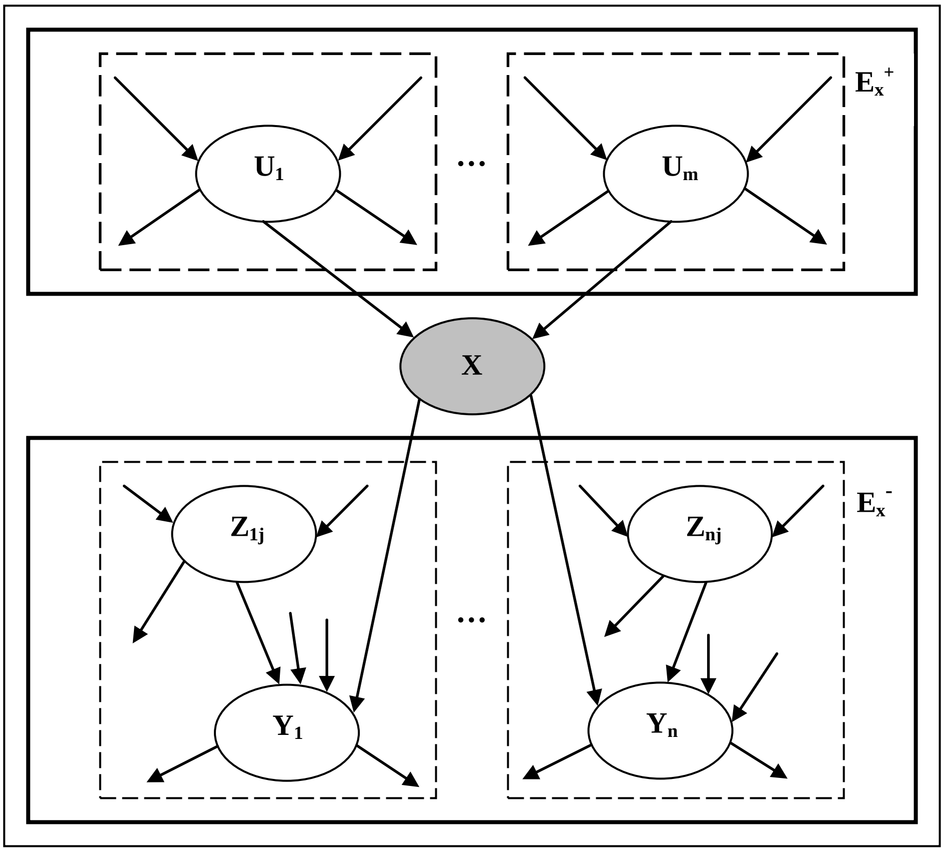

Fig. 5.4

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

$$E_X^+ = \bigcup_{i=1}^{m} E_{U_i\setminus X} \qquad (1)$$

În mod similar, $E_{Y_i\setminus X}$ semnifică mulţimea tuturor dovezilor conectate la Y_i prin intermediul părinţilor săi, *cu excepţia* lui X. Mulţimea E_X^- poate fi deci partiţionată în mulţimile $E_{Y_1\setminus X},\ldots,E_{Y_n\setminus X}$ şi este de forma

$$E_X^- = \bigcup_{i=1}^{n} E_{Y_i\setminus X} \qquad (2)$$

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

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

- Exprimă $P(X\,|\,E)$ în termenii contribuţiilor lui E_X^+ şi E_X^-.
- Calculează contribuţia mulţimii E_X^+ calculând efectul ei asupra părinţilor lui X şi apoi transmiţând acest efect lui X. (Calcularea efectului asupra fiecărui părinte al lui X este o secvenţă recursivă a problemei calculării efectului asupra lui X).
- Calculează contribuţia mulţimii E_X^- calculând efectul ei asupra fiilor lui X şi apoi transmiţând acest efect lui X. (Calcularea

efectului asupra fiecărui fiu al lui X reprezintă o secvenţă 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

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

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

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

Întrucât X d-separă în cadrul reţelei pe E_X^+ de E_X^-, putem folosi independenţa condiţionată pentru a simplifica al doilea termen al numărătorului. De asemenea, putem trata $1/P\left(E_X^- \mid E_X^+\right)$ ca pe o constantă de normalizare, obţinând:

$$P\left(X \mid E\right)=\alpha\, P\left(X \mid E_X^+\right)P\left(E_X^- \mid X\right) \tag{3}$$

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

Vom calcula $P\left(X \mid E_X^+\right)$ luând în consideraţie toate configuraţiile posibile ale părinţilor lui X, precum şi cât de probabile sunt acestea fiind dată mulţimea E_X^+. În cazul fiecărei configuraţii date, probabilitatea lui X se cunoaşte direct din tabelul probabilităţilor condiţionate.

Fie $\mathbf{U}$ vectorul părinţilor $U_1,\ldots,U_m$ 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^+. Ţinând cont de acest fapt şi de formula (1), obţinem egalitatea:

$$P\left(\{U_i = u_i\} \mid E_X^+\right) = P\left(\{U_i = u_i\} \mid E_{U_i\setminus X}\right) \quad i = \overline{1,m} \tag{4}$$

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

$$P\left(X \mid E_X^+\right) = P\left(X \bigcap \bigcup_{(u_1,\ldots,u_m)} \{U_1 = u_1,\ldots,U_m = u_m\} \mid E_X^+\right) \tag{5}$$

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

$$P\left(X \mid E_X^+\right) = \sum_{(u_1,\ldots,u_m)} P\left(X \bigcap \{U_1 = u_1,\ldots,U_m = u_m\} \mid E_X^+\right) \tag{6}$$

Aplicând, în continuare, versiunea condiţionată a formulei de înmulţire a probabilităţilor membrului drept din (6), obţinem:

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

$$\cdot P\left(X \mid \{U_1 = u_1,\ldots,U_m = u_m\}, E_X^+\right) \tag{7}$$

---

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

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 ţine cont de independenţa 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 partiţionat în $E_{U_1 \setminus X},\ldots,E_{U_m \setminus X}$. Întrucât $U_1,\ldots,U_m$ sunt independente,

$$P\left(\left\{U_1=u_1,\ldots,U_m=u_m\right\} \mid E_X^+\right)=\prod_{i=1}^{m} P\left(\left\{U_i=u_i\right\} \mid E_X^+\right) \tag{8}$$

şi, ţinând cont de (1), avem:

$$\prod_{i=1}^{m} P\left(\left\{U_i=u_i\right\} \mid E_X^+\right)=\prod_{i=1}^{m} P\left(\left\{U_i=u_i\right\} \mid E_{U_i \setminus X}\right) \tag{9}$$

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

$$P\left(X \mid \left\{U_1=u_1,\ldots,U_m=u_m\right\}, E_X^+\right)=P\left(X \mid \left\{U_1=u_1,\ldots,U_m=u_m\right\}\right) \tag{10}$$

Ţinând cont de (7), (8), (9) şi (10), obţinem:

$$P\left(X \mid E_X^+\right)=\sum_{(u_1,\ldots,u_m)} P\left(X \mid \left\{U_1=u_1,\ldots,U_m=u_m\right\}\right)\prod_{i=1}^{m} P\left(\left\{U_i=u_i\right\} \mid E_{U_i \setminus X}\right) \tag{11}$$

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

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

Ecuaţia (12) sugerează deja un algoritm. Astfel, $P\left(X \mid \mathbf{u}\right)$ este repartiţia lui X condiţionată de realizarea $\left(U_1=u_1,\ldots,U_m=u_m\right)$. Valoarea acestei probabilităţi poate fi luată din tabelul de probabilităţi condiţionate asociat lui X. Calculul fiecărei probabilităţi $P\left(\left\{U_i = u_i\right\} \mid E_{U_i \backslash X}\right)$ reprezintă o secvenţă recursivă a problemei iniţiale, aceea de a calcula $P\left(X \mid E\right)$, adică $P\left(X \mid E_{X\backslash}\right)$. Vom mai nota aici faptul că mulţimea variabilelor dovezi care intervin în apelarea recursivă reprezintă o submulţime a celor din apelarea iniţială, ceea ce constituie un indiciu că procedura de calcul se va termina într-un număr finit de paşi, ea reprezentând într-adevăr un algoritm.

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. – **CURSUL URMATOR, la sfarsitul caruia vom putea formula un algoritm pentru raspunsul la interogari.**