Laboratorul 5: **Rețele Bayesiene**

Creierul uman are o capacitate remarcabilă de a lua decizii în prezența informațiilor incerte sau incomplete. Facem predicții, uneori inconștient, despre evenimente din viitor sau pe care nu le putem observa direct și asta ne ajută să interacționăm atât de eficient cu lumea din jur.

Este necesar, deci, atunci când discutăm despre inteligență artificială să ne punem problema reprezentării incertitudinii unor evenimente, fie că aceasta provine din absența unor observații, fie că este consecința unor procese aleatoare, nedeterministe (e.g. aruncarea unui zar). Cadrul matematic care oferă conceptele utile acestei discuții îl reprezintă teoria probabilităților. În cele ce urmează sunt reluate câteva noțiuni elementare necesare.

1 Teoria probabilităților

1.1 Probabilitatea unui eveniment

Fie `X` o variabilă aleatoare care poate lua valoarea 0 sau 1 (`X` este o variabilă aleatoare discretă binară). Definim *probabilitatea* producerii evenimentului $\texttt{X} = 1$ ca fiind raportul dintre numărul de dăți în care variabila `X` ia valoarea 1 și numărul total de experimente. Vom nota această probabilitate astfel: $p\left(\texttt{X} = 1\right)$. Prin definiție orice probabilitate este un număr între 0 și 1.

Pentru simplificare vom nota în continuare cu $p\left(X\right)$ probabilitatea unui eveniment oarecare X.

Dacă $p\left(X\right) = 1$ atunci X este un eveniment sigur (e.g. $\texttt{X} = 0 \vee \texttt{X} = 1$).

Dacă $p\left(X\right) = 0$ atunci X este un eveniment imposibil.

Vom nota cu $\neg X$ evenimentul contrar (sau complementar) evenimentului X.

$$p\left(X\right) = 1 - p\left(\neg X\right) \tag{1}$$

1.2 Evenimente independente. Probabilități condiționate

Evenimente independente

Două evenimente X și Y se numesc *independente* dacă probabilitatea evenimentului combinat $X \cap Y$ este produsul probabilităților individuale de producere a celor două evenimente (vezi Formula 2). Vom nota fapul că două evenimente sunt independente astfel: $X \perp Y$. O observație importantă este aceea că dacă

N evenimente sunt independente două câte două, acestea nu sunt necesar independente în ansamblu.

$$p(X \cap Y) = p(X) \cdot p(Y) \tag{2}$$

Două evenimente X și Y se numesc *incompatibile* dacă acestea nu se pot produce simultan: $p(X \cap Y) = 0$.

Probabilități condiționate

Fie X și Y două evenimente și $p(Y) \neq 0$. Probabilitatea lui X condiționat de Y este dată de raportul dintre probabilitatea evenimentului combinat și probabilitatea evenimentului Y (vezi Formula 3). Intuitiv, aceasta descrie șansele de a se produce evenimentul X atunci când evenimentul Y s-a realizat.

$$p(X|Y) = \frac{p(X \cap Y)}{p(Y)} \tag{3}$$

Formula 3 poate fi aplicată succesiv pentru a *desfășura* probabilitatea unui eveniment combinat din N evenimente (vezi Formula 4).

$$\begin{aligned} p(X_1 \cap X_2 \cap \ldots \cap X_N) &= p(X_1)\, p(X_2 \cap \ldots \cap X_N | X_1) \\ &= p(X_1)\, p(X_2|X_1)\, p(X_3 \cap \ldots \cap X_N | X_1 \cap X_2) \\ &\ldots \\ &= p(X_1)\, p(X_2|X_1) \ldots p(X_N | X_1 \cap X_2 \cap \ldots \cap X_{N-1}) \end{aligned} \tag{4}$$

Dacă un eveniment sigur Y se poate descompune în N evenimente incompatibile $Y_1, Y_2, \ldots, Y_N$, atunci pentru orice eveniment X este adevărată identitatea 5 (forumla probabilității totale).

$$p(X) = \sum_{n=1}^{N} p(X|Y_n)\, p(Y_n) = \sum_{n=1}^{N} p(X \cap Y_n) \tag{5}$$

O aplicație uzuală a acestei reguli este descompunerea evenimentului sigur în Y și $\neg Y$ (Formula 6).

$$p(X) = p(X|Y)\, p(Y) + p(X|\neg Y)\, p(\neg Y) \tag{6}$$

Exemplu Să presupunem că un agent vrea să își evalueze șansele de a câștiga un joc cu zar. Fie C evenimentul în care agentul câștigă jocul. Presupunând că agentul cunoaște probabilitatea ca zarul să pice pe fiecare dintre cele șase fețe ale sale $p(\mathtt{Z} = i), \forall i \in \{1, \ldots, 6\}$, acesta trebuie să evalueze apoi care sunt șansele de a câștiga în fiecare dintre acele cazuri $p(C|\mathtt{Z} = i)$. Atunci, șansele ca agentul să câștige jocul sunt date de cantitea următoare (Formula 7):

$$p(C) = \sum_{i=1}^{6} p(C|\mathtt{Z} = i) \cdot p(\mathtt{Z} = i) \tag{7}$$

1.3 Teorema lui Bayes

Pentru două evenimente X și Y, aplicând Formula 5 și ținând cont de simetria $p(X \cap Y) = p(Y \cap X)$ se obține *Teorema lui Bayes*:

$$p(X|Y) = \frac{p(Y|X)\, p(X)}{p(Y)} \tag{8}$$

Teorema lui Bayes paote fi interpretată ca o condiționare cauzală. Frumusețea acestei formule vine din faptul că exprimă probabilitatea unei cauze pe baza unei observații făcute, întorcând direcția de condiționare. Probabilitatea posterioară $p(X|Y)$ este exprimată ca produsul dintre verosimilitatea observației făcute dată fiind cauza X, $p(Y|X)$, și probabilitatea apriorică a cauzei $p(X)$ împărțit la probabilitatea totală a evenimentului Y.

O aplicație uzuală a Teoremei lui Bayes este aplicarea în cazul evenimentelor complementare precum în Formula 9.

$$p(X|Y) = \frac{p(Y|X)\, p(X)}{p(Y|X)\, p(X) + p(Y|\neg X)\, p(\neg X)} \tag{9}$$

Exemplu

Se consideră o anume formă de gripa ce afectează un procent cunoscut din populație (1%) și un test ce semnalează cu o anumită certitudine prezența acestei boli. Mai precis, se cunosc următoarele probabilități:

- $p(Gripa) = 0.01$ - incidența acestei forme de gripa ete 1%;
- $p(+|Gripa) = 0.9$ - probabilitatea ca testul să indice corect prezența gripei în cazul unui pacient bolnav;
- $p(+|\neg Gripa) = 0.2$ - probabilitatea ca testul să indice greșit prezența gripei în cazul unui pacient sănătos.

Să notăm și probabilitățile evenimentelor complementare:

- $p(\neg Gripa) = 0.99$ - probabilitatea evenimentului complementar (indivizii sănătoși din rândul populației);
- $p(-|Gripa) = 0.1$ - probabilitatea ca testul să NU indice prezența gripei în cazul unui pacient bolnav (negativ fals);
- $p(-|\neg Gripa) = 0.8$ - probabilitatea ca testul să iasă negativ în cazul unui individ sănătos.

Aplicând teorema lui Bayes putem afla care este probabilitatea ca un individ să fie bolnav dacă testul a ieșit pozitiv:

$$\begin{aligned} p\left(Gripa|+\right) &= \frac{p\left(+|Gripa\right)p\left(Gripa\right)}{p\left(+\right)} \\ &= \frac{p\left(+|Gripa\right)p\left(Gripa\right)}{p\left(+|Gripa\right)p\left(Gripa\right)+p\left(+|\neg Gripa\right)p\left(\neg Gripa\right)} \\ &= \frac{0.9\cdot 0.01}{0.9+0.01+0.2\cdot 0.99} \approx 0.04348 \end{aligned} \tag{10}$$

Artificiu de calcul Uneori, în teorema lui Bayes numitorul este greu de calculat. Se poate evita acest calcul direct folosindu-se evenimentul complementar.

$$p\left(X|Y\right) = \frac{p\left(Y|X\right)p\left(X\right)}{p\left(Y\right)} \overset{\text{not.}}{=} \frac{P1}{\gamma} \tag{11}$$

$$p\left(\neg X|Y\right) = \frac{p\left(Y|\neg X\right)p\left(\neg X\right)}{p\left(Y\right)} \overset{\text{not.}}{=} \frac{P2}{\gamma} \tag{12}$$

Calculându-se doar numărătorii din Ecuațiile 11 și 12, se rezolvă apoi următorul sistem de ecuații:

$$\gamma P1 + \gamma P2 = 1 \tag{13}$$

$$p\left(X|Y\right) = \gamma P1 \tag{14}$$

2 Rețele Bayesiene

Rețelele Bayesiene sunt modele grafice utilizate pentru reprezentarea eficientă a unor distribuții comune peste un spațiu mai mare de variabile aleatoare. Rețelele Bayesiene sunt grafuri orientate aciclice în care nodurile reprezintă variabile aleatoare, iar arcele reprezintă relații cauzale directe între acestea.

Rețele Bayesiene pot lucra și cu variabile continue, dar în cadrul acestei expuneri ne vom referi doar la rețele ce descriu probabilități pentru variabile discrete.

Figura 1 reprezintă o rețea bayesiană cu două variabile binare: `A` și `B`. Pentru a simplifica notațiile vom considera evenimentele $A \equiv \texttt{A} = 1$ și $\neg A \equiv \texttt{A} = 0$ notând, deci, $p\left(A\right) = p\left(\texttt{A} = 1\right)$ și probabilitatea evenimentului complementar $p\left(\neg A\right) = p\left(\texttt{A} = 0\right)$. Convenția este folosită pentru toate variabilele binare `B`, `C`, `D`, etc. ce vor apărea mai jos.

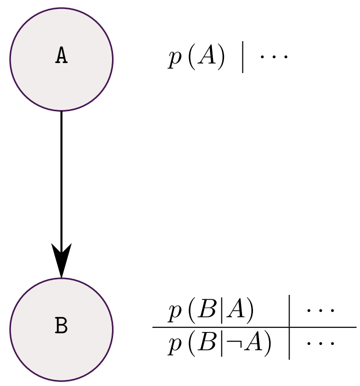


Figura 1: Rețea Bayesiană simplă

Fiecare nod dintr-o rețea bayesiană este însoțit de o tabelă de probabilități condiționate care descriu șansele ca variabila respectivă să ia o valoare atunci când variabilele părinte (cauzele) au fost observate. Pentru o variabilă care nu are părinți (cauze), probabilitățile de realizare ale acesteia nu vor fi condiționate.

Mai precis, pentru rețeaua bayesiană din Figura 1 trebuie precizate $p(A)$, $p(B|A)$ și $p(B|\neg A)$.

În cazul general, în care fiecare variabilă $\mathtt{X}$ poate lua valori dintr-o mulțime de valori $\mathcal{D}(\mathtt{X})$, numărul de valori ce trebuie specificate pentru un nod oarecare este $\left(|\mathcal{D}(\mathtt{X})| - 1\right) \prod_{\mathtt{Y} \in Par(\mathtt{X})} |\mathcal{D}(\mathtt{Y})|$, unde $|M|$ este cardinalul mulțimii M, iar $Par(\mathtt{X})$ reprezintă mulțimea părinților (cauzelor) variabilei $\mathtt{X}$.

Întrebarea 1 Fie mulțimea de variabile aleatoare binare $\{\mathtt{A}, \mathtt{B}, \mathtt{C}, \mathtt{D}, \mathtt{E}, \mathtt{F}, \mathtt{G}\}$. Câte valori ar trebui specificate (sau învățate într-un proces de antrenare) pentru a descrie probabilitatea comună $p(\mathtt{A}, \mathtt{B}, \mathtt{C}, \mathtt{D}, \mathtt{E}, \mathtt{F}, \mathtt{G})$? Câte valori trebuie specificate pentru a descrie rețeaua bayesiană din Figura 2? Răspundeți la aceleași două întrebări considerând că variabilele sunt ternare (pot lua trei valori).

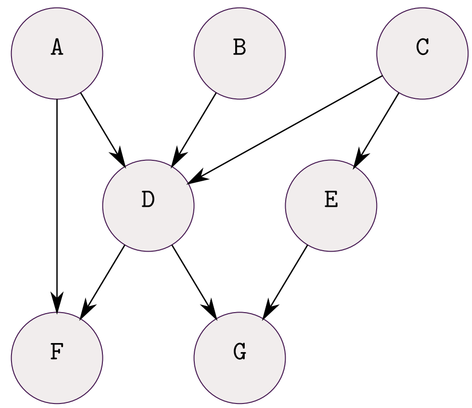


Figura 2: Rețea Bayesiană pentru întrebarea 1

2.1 Independență condițională

Unul dintre conceptele cheie din teoria probabilităților care face ca lucrul cu rețelele bayesiane să fie eficient este *independența condițională*. Dacă observarea unei variabile $\mathtt{Y}$ face ca variabila $\mathtt{X}$ să nu depindă de o a treia variabilă $\mathtt{Z}$ atunci spunem că $\mathtt{X}$ este condițional independentă de $\mathtt{Z}$ dată fiind $\mathtt{Y}$. Intuitiv, această relație indică faptul că atunci când rezultatul lui $\mathtt{Y}$ este cunoscut, observarea lui $\mathtt{Z}$ nu ar aduce nicio informație suplimentară despre $\mathtt{X}$ (sau, echivalent, nu ar modifica (in)certitudinea asupra variabilei $\mathtt{X}$). Una dintre notațiile uzuale pentru independența condițională este cea din Formula 15.

$$\mathtt{X} \perp \mathtt{Z} | \mathtt{Y} \iff p\left(\mathtt{X}|\mathtt{Y}, \mathtt{Z}\right) = p\left(\mathtt{X}|\mathtt{Y}\right) \tag{15}$$

Relația de independență condițională se generalizează și la mulțimi de variabile. În Formula 16, $\mathbf{X}$, $\mathbf{Y}$ și $\mathbf{Z}$ reprezintă mulțimi de variabile aleatoare.

$$\mathbf{X} \perp \mathbf{Z} | \mathbf{Y} \iff p\left(\mathbf{X}|\mathbf{Y}, \mathbf{Z}\right) = p\left(\mathbf{X}|\mathbf{Y}\right) \tag{16}$$

Pentru a putea investiga aplicațiile independenței condiționale în rețele bayesiene, mai jos sunt descrise trei posibile relații în care se pot găsi trei variabile aleatoare.

Lanț de cauzalitate Fie rețeaua bayesiană din Figura 3 cu trei variabile $\mathtt{A}$, $\mathtt{B}$ și $\mathtt{C}$. Să presupunem că aceasta modelează următoarea situație: variabila $\mathtt{A}$ reprezintă faptul că este iarnă, variabila $\mathtt{B}$ starea de a ninge afară, iar variabila $\mathtt{C}$ faptul că Ionuț se dă cu sania. În primul caz să presupunem că nu știm dacă afară ninge sau nu. Informația despre anotimpul curent (dacă este iarnă sau nu, adică observarea variabilei $\mathtt{A}$) modifică ceea ce credem despre faptul că Ionuț se dă cu sania. Mai precis, într-un alt anotimp decât iarna, acest lucru ar părea improbabil, deoarce este foarte puțin probabil să ningă. Iarna, în schimb, putem crede că sunt șanse mari să ningă, drept urmare sunt șanse mai mari ca

Ionuț să se dea cu sania. Dar, odată ce aflăm dacă ninge sau nu, informația despre anotimpul curent nu mai poate influența șansele pe care le acordăm evenimentului Ionuț se dă cu sania. În acest caz, deci, $\mathtt{A} \perp \mathtt{C}|\mathtt{B}$ și $\mathtt{A} \not\perp \mathtt{C}$.

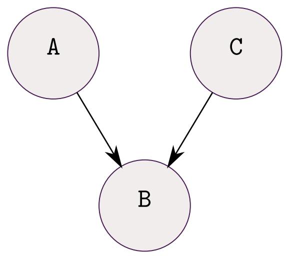
(a) $\mathtt{A} \not\perp \mathtt{C}$ (b) $\mathtt{A} \perp \mathtt{C}|\mathtt{B}$

Figura 3: Variabilele $\mathtt{A}$ și $\mathtt{B}$ devin condițional independente atunci când variabila $\mathtt{C}$ este observată (imaginea b).

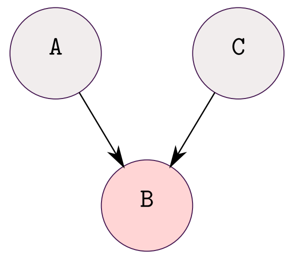
Cauză comună Fie rețeaua bayesiană din Figura 4. În acest caz variabila $\mathtt{B}$ condiționează atât variabila $\mathtt{A}$, cât și variabila $\mathtt{C}$. Să asociem trei evenimente concrete celor trei variabile: evenimentul *plouă* este descris de variabila binară $\mathtt{B}$, variabila $\mathtt{A}$ reprezintă faptul că orașul București este aglomerat, iar variabila $\mathtt{C}$ faptul că Ionuț poartă o umbrelă. Exact ca în cazul precedent, să presupunem întâi că nu știm dacă afară plouă. Faptul că îl vedem pe Ionuț cu o umbrelă ne poate face mai încrezători în faptul că traficul este aglomerat deoarece este foarte probabil să plouă. Totuși, observând dacă afară plouă sau nu, umbrela lui Ionuț nu ne mai aduce nicio informație despre traficul din București. Deci, în acest caz $\mathtt{A} \perp \mathtt{C}|\mathtt{B}$ și $\mathtt{A} \not\perp \mathtt{C}$.

(a) $\mathtt{A} \not\perp \mathtt{C}$ (b) $\mathtt{A} \perp \mathtt{C}|\mathtt{B}$

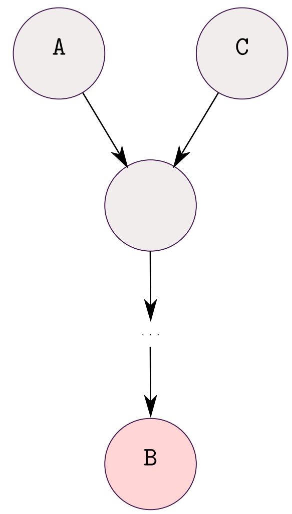
Figura 4: Variabilele $\mathtt{A}$ și $\mathtt{B}$ devin condițional independente atunci când variabila $\mathtt{C}$ este observată (imaginea b).



(a) $\mathtt{A} \perp \mathtt{C}$



(b) $\mathtt{A} \not\perp \mathtt{C} | \mathtt{B}$



(c) $\mathtt{A} \not\perp \mathtt{C} | \mathtt{B}$

Figura 5: Variabilele `A` și `B` sunt condițional independente atunci când variabila `C` nu este observată (imaginea a), dar depind una de cealaltă după observarea unui copil comun (imaginea b) sau a unui descendent al acestuia (imaginea c).

Explaining away Fie rețeaua bayesiană din Figura 5. În acest caz: $\mathtt{A} \not\perp \mathtt{C} | \mathtt{B}$ și $\mathtt{A} \perp \mathtt{C}$. Să încercăm să suprapunem un exemplu. Ionuț este microbist și echipa lui preferată este Barcelona care are un meci important cu rivala din Madrid. De asemenea, Ionuț este programator și așteaptă verdictul unui interviu pe care l-a avut de curând. Să zicem că variabila `A` corespunde evenimentului *Barcelona a câștigat meciul cu Real Madrid*, iar variabila `C` reprezintă evenimentul *Ionuț a luat interviul*. Fără nicio informație suplimentară cele două evenimente nu au nicio legătură și informații despre unul dintre ele nu ne pot schimba probabilitatea pe care o asociem celuilalt. Dacă, în schimb, observăm variabila `B` și aflăm că Ionuț este fericit, atunci apare nevoia unei explicații. De ce este Ionuț fericit? În cazul acesta, dacă aflăm că rezultatul meciului a fost negativ, din nevoia de oferi o explicație pentru bucuria lui Ionuț, vom avea o certitudine ridicată că acesta a luat interviul. Variabilele de care depinde `B` devin dependente deoarece se activează un raționament intercauzal în căutarea unei explicații.

2.2 D-Separabilitate

Noțiunea de *D-separabilitate* folosește o analiză a grafului care descrie rețeaua bayesiană pentru a determina dacă două mulțimi de variabile sunt independente condițional sau nu.

O *cale neorientată* între o mulțime de noduri $\mathbf{X}$ și o mulțime de noduri $\mathbf{Z}$ este orice secvență de arce (indiferent de orientarea lor) care pleacă de la un membru al mulțimii $\mathbf{X}$ și ajunge la un membru al mulțimii $\mathbf{Z}$.

Dându-se o mulțime de noduri **O** observate, o cale **p** este *blocată* dacă una dintre următoarele trei condiții este adevărată:

1. există un nod $\mathtt{Y} \in \mathbf{O}$ care reprezintă destinația unui arc din **p** și sursa unui alt arc din **p** (înlănțuire cauzală);
2. există un nod $\mathtt{Y} \in \mathbf{O}$ care reprezintă sursa a două arce diferite din **p** (cauză comună);
3. niciun nod `Y` care reprezintă destinația a două arce din **p** nu se află în **O** și nici vreun descendent de-ai lui nu se află în **O**.

O mulțime de noduri **O** *d-separă* două mulțimi **X** și **Z** dacă orice cală neorientată între **X** și **Z** este blocată prin observarea variabilelor din **O**.

3 Exemplu și probleme

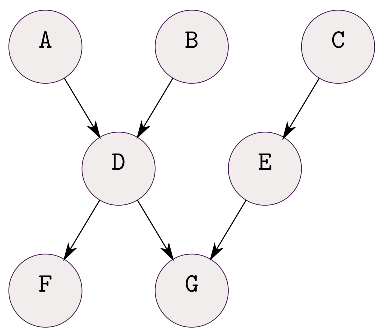


Figura 6: Rețea bayesiană întrebarea 2.

Fie rețeaua bayesiană din Figura 6.

Întrebarea 2. De câți parametri este nevoie pentru a descrie complet rețeaua bayesiană din Figura 6? De câți ar fi fost nevoie într-o reprezentare tabulară completă a $p\left(\mathtt{A}, \mathtt{B}, \mathtt{C}, \mathtt{D}, \mathtt{E}, \mathtt{F}, \mathtt{G}\right)$?

Întrebarea 3. Precizați care dintre următoarele relații de independență condițională sunt adevărate în rețeaua din Figura 6. Analizați folosind definiția pentru *d-separabilitate*.

a) $\mathtt{A} \perp \mathtt{G} | \mathtt{D}$

b) $\mathtt{A} \perp \mathtt{B} | \mathtt{F}$

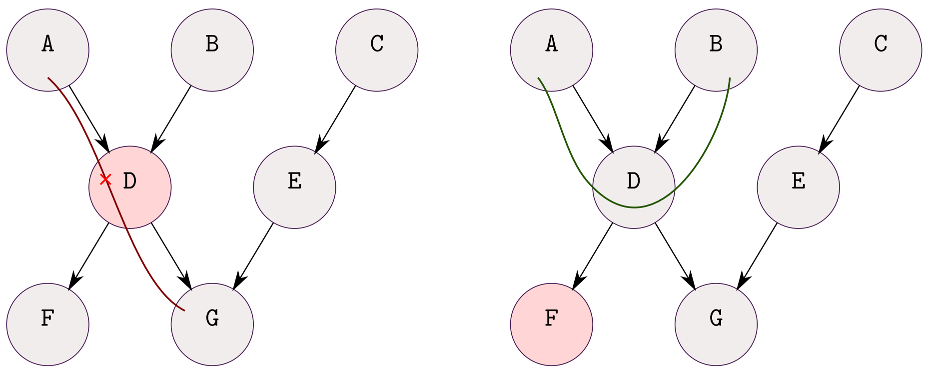
c) $\mathtt{A} \perp \mathtt{C}$

d) $\texttt{A} \perp \texttt{C}|\texttt{E}$

e) $\texttt{F} \perp \texttt{G}|\texttt{A}$

f) $\texttt{B} \perp \texttt{C}|\texttt{G}$

g) $\texttt{F} \perp \texttt{C}|\texttt{G}$



(a) Întrebarea 3.a. (b) Întrebarea 3.b.

Figura 7: În imaginea a variabilele `A` și `G` sunt condițional independente prin observarea variabilei `D` deoarece singura cale dintre acestea este blocată. În imaginea b variabilele `A` și `B` sunt dependente deoarece observarea unui descendent al variabilei `D` activează calea $\texttt{A} - \texttt{D} - \texttt{B}$.

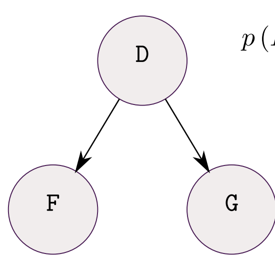
Întrebarea 4. Considerând doar variabilele `D`, `F`, `G`, precum în Figura 8, calculați următoarele probabilități:

a) $p\left(D|F \cap G\right)$

b) $p\left(F|G\right)$

$p\left(D\right)$	0.7

$p\left(F\|D\right)$	0.8
$p\left(F\|\neg D\right)$	0.5



$p\left(G\|D\right)$	0.25
$p\left(G\|\neg D\right)$	0.65

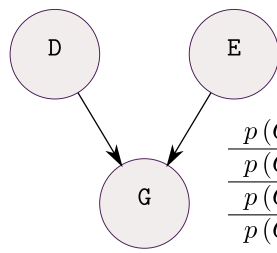
Figura 8: Rețea bayesiană pentru întrebarea 4.

Întrebarea 5. Considerând doar variabilele `D`, `E`, `G`, precum în Figura 9, calculați următoarele probabilități :

a) $p\left(D|E\right)$

b) $p\left(D|G \cap E\right)$

$p\left(D\right)$	0.7



$p\left(E\right)$	0.25

$p\left(G\|D \cap E\right)$	0.9
$p\left(G\|D \cap \neg E\right)$	0.8
$p\left(G\|\neg D \cap E\right)$	0.6
$p\left(G\|\neg D \cap \neg E\right)$	0.2

Figura 9: Rețea bayesiană pentru întrebarea 5.

Indiciu pentru întrebarea 5.b Putem descompune $p\left(\mathtt{X}, \mathtt{Y}, \mathtt{Z}\right)$ folosind Formula 4 în două moduri:

$$p\left(\mathtt{X}, \mathtt{Y}, \mathtt{Z}\right) = p\left(\mathtt{X}\right) p\left(\mathtt{Y}|\mathtt{X}\right) p\left(\mathtt{Z}|\mathtt{Y} \cap \mathtt{X}\right) \tag{17}$$

$$p\left(\mathtt{X}, \mathtt{Y}, \mathtt{Z}\right) = p\left(\mathtt{X}\right) p\left(\mathtt{Z}|\mathtt{X}\right) p\left(\mathtt{Y}|\mathtt{Z} \cap \mathtt{X}\right) \tag{18}$$

Din Formulele 17 și 18 se obține următoarea egalitate (o expresie similară Formulei lui Bayes în care toate probabilitățile sunt condiționate de o altă variabilă):

$$p\left(\mathtt{Y}|\mathtt{X}\right) p\left(\mathtt{Z}|\mathtt{Y} \cap \mathtt{X}\right) = p\left(\mathtt{Z}|\mathtt{X}\right) p\left(\mathtt{Y}|\mathtt{Z} \cap \mathtt{X}\right) \tag{19}$$

Întrebarea 6. Urmărind rețeaua bayesiană din Figura 10, calculați următoarele probabilități:

a) $p\left(B|F\right)$

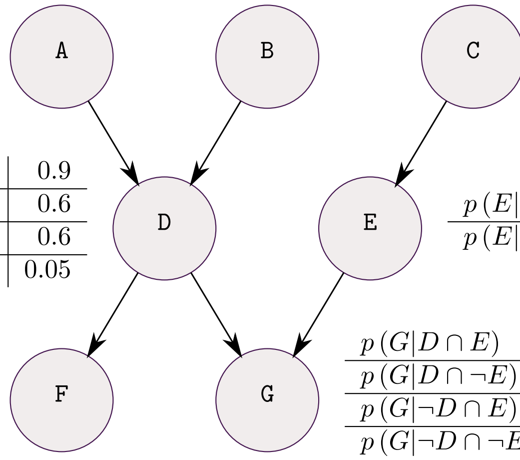
b) $p\left(A|\neg F \cap \neg G\right)$

c) $p\left(B|\neg C\right)$

$p(B)$	0.8

$p(A)$	0.6

$p(C)$	0.1



$p(D \mid A \cap B)$	0.9
$p(D \mid A \cap \neg B)$	0.6
$p(D \mid \neg A \cap B)$	0.6
$p(D \mid \neg A \cap \neg B)$	0.05

$p(E \mid C)$	0.7
$p(E \mid \neg C)$	0.2

$p(F \mid D)$	0.8
$p(F \mid \neg D)$	0.5

$p(G \mid D \cap E)$	0.9
$p(G \mid D \cap \neg E)$	0.8
$p(G \mid \neg D \cap E)$	0.6
$p(G \mid \neg D \cap \neg E)$	0.2

Figura 10: Rețea bayesiană cu 7 noduri pentru întrebarea 6.