# Alkalmazott mesterséges intelligencia (AMI)

http://www.mit.bme.hu/oktatas/targyak/vimibb01

4. ea. (2023. ősz)

# Valószínűségi (Bayes) hálók

http://mialmanach.mit.bme.hu/aima/ch14

Jegyzet 14. fejezet


https://www.esrcheck.com/2023/06/05/artificial-intelligence-ai-experts-sign-statement-on-ai-risk/

Előadó: Pataki Béla

a fóliák
Dobrowiecki Tadeusz és
Hullám Gábor anyagainak
felhasználásával készültek

BME I.E. 414, 463-26-79
pataki@mit.bme.hu,
http://www.mit.bme.hu/general/staff/pataki

# Eddigiekben:

+ a világ állapotátmeneti modellje (Arad->Nagyszeben stb.)
+ logikai modellje (állítások, ítéletek, logikai következtetés)
+ következő: események, bekövetkezési valószínűséggel (véletlen változók), ok-okozati összefüggések

# Együttes valószínűség-eloszlás: $P(x_1, x_2, \dots, x_N)$

Példa: Annak valószínűsége, hogy megverjük futballban Olaszországot, miközben **(=és)** a benzinár 100 Ft/liter **és** szivárvány van az égen *P*(EngvsHun=HunWin, BenzinÁr=100, Szivárvány=Van) = $10^{-8}$.

Másképpen jelölve: *P*(HunWin, Benzin100, Szivárvány) = $10^{-8}$ .

(gyk. nem a szivárvány kis esélye miatt ilyen kicsi…. ☺)

*P*(HunWin, Benzin100, $\neg$Szivárvány) = $10^{-6}$

*P*(HunWin, $\neg$ Benzin100, $\neg$Szivárvány) = $17 \cdot 10^{-2}$ (=17%)

**Jó hír**: az együttes eloszlás birtokában minden olyan kérdésre kapunk választ, amely a benne szereplő véletlen változók viszonyára és tulajdonságaira vonatkozik.

$P(x_1, x_2, \ldots, x_N)$ esetén kell **$2^N - 1$** független valószínűségérték (bináris változók esetén). Tehát a válaszhoz a változók számával **exponenciális** növekedést mutató számú valószínűség ismerete lehet szükséges. (Azért va a végén -1, mert a $2^N$ lehetőség valószínűségének összege 1, tehát az „utolsó” kiszámítható)


**Rossz hír**: nemigen megy 10-nél sokkal több változót tartalmazó eloszlások táblázatos megadása

# Egy többé-kevésbé életszerű (egyszerűsített) példa

(a Russell-Norvig könyvből)

A lakásunkban van egy riasztó, ami betörésnél riaszt.

Sajnos nem tökéletes: betörés esetén csak 95% valószínűséggel riaszt, de ha földrengés van, akkor is riaszt (29% valószínűséggel). Sőt, ha egyik sincs, akkor is riaszt 0,1% valószínűséggel. Ha a földrengés esetén van betörés is, akkor is 95%, hogy megszólal.

Két szomszédot, akik mindig otthon ülnek, Jánost és Máriát kértünk meg, hogy telefonáljanak a munkahelyünkre, ha riasztást hallanak.

Persze ők se hibátlanok: János 90% eséllyel hallja meg a riasztást, míg Mária csak 70% eséllyel. Viszont néha más zajokat is riasztásnak gondolnak, és telefonálnak akkor is, amikor nincs riasztás. Ez Jánosnál 5% eséllyel fordul elő, Máriánál 1% eséllyel.

1. Mekkora az esélye, hogy nem telefonál egyik sem, ha betörés van?

2. Mekkora az esélye, hogy legalább az egyikük telefonál, ha betörés van?

3. Ha mindketten telefonálnak, mekkora az esélye, hogy nincs betörés?

4. …és még sok hasonló kérdést feltehetünk….

# Bayes (valószínűségi) hálók

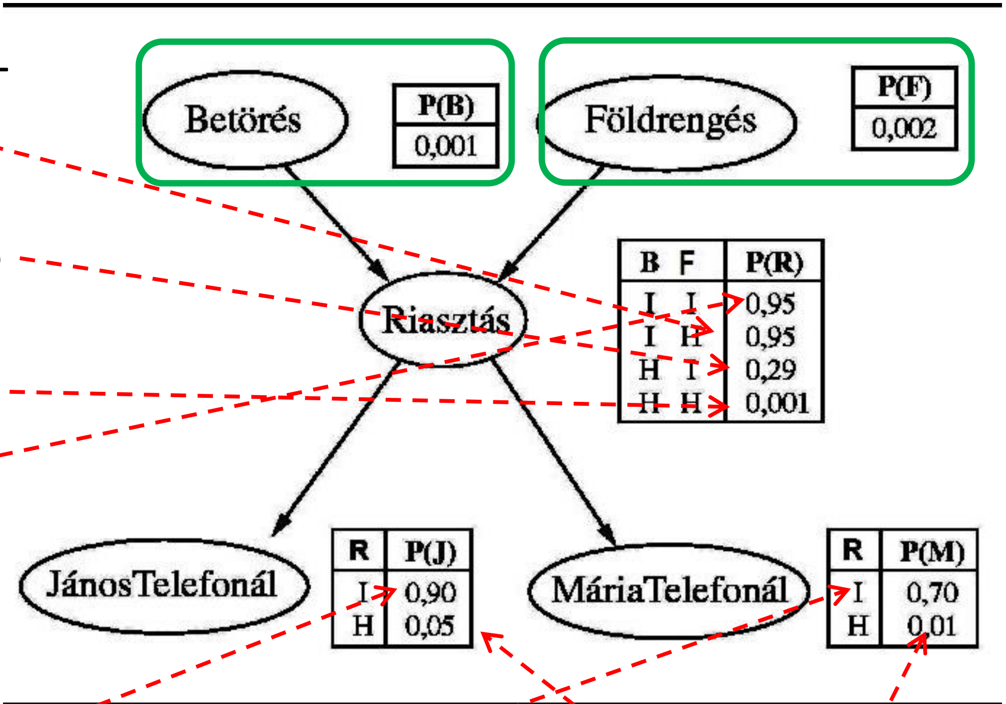
Az együttes eloszlást $P(x_1, x_2, \ldots, x_N)$ -t adjuk meg, de kihasználjuk a változók közt fennálló feltételes függetlenségeket!

Betörés esetén csak 95% valószínűség-gel riaszt **(NEM FÜGG A TELEFONÁLÁS-TÓL!)**

Ha földrengés van, akkor is riaszt (29% valószínűséggel)

Ha egyik sincs, akkor is riaszt 0,1% valószínűséggel.

Ha a földrengés esetén van betörés is, akkor is 95%, hogy megszólal.




Jánost és Máriát kértünk meg, hogy telefonáljanak a munkahelyünkre, **ha riasztást hallanak**. **(NEM KÖZVETLENÜL A BETÖRÉSRE VAGY A FÖLDRENGÉSRE REAGÁLNAK? HANEM A RISZTÁSRA !!!)**

János 90% eséllyel hallja meg a riasztást

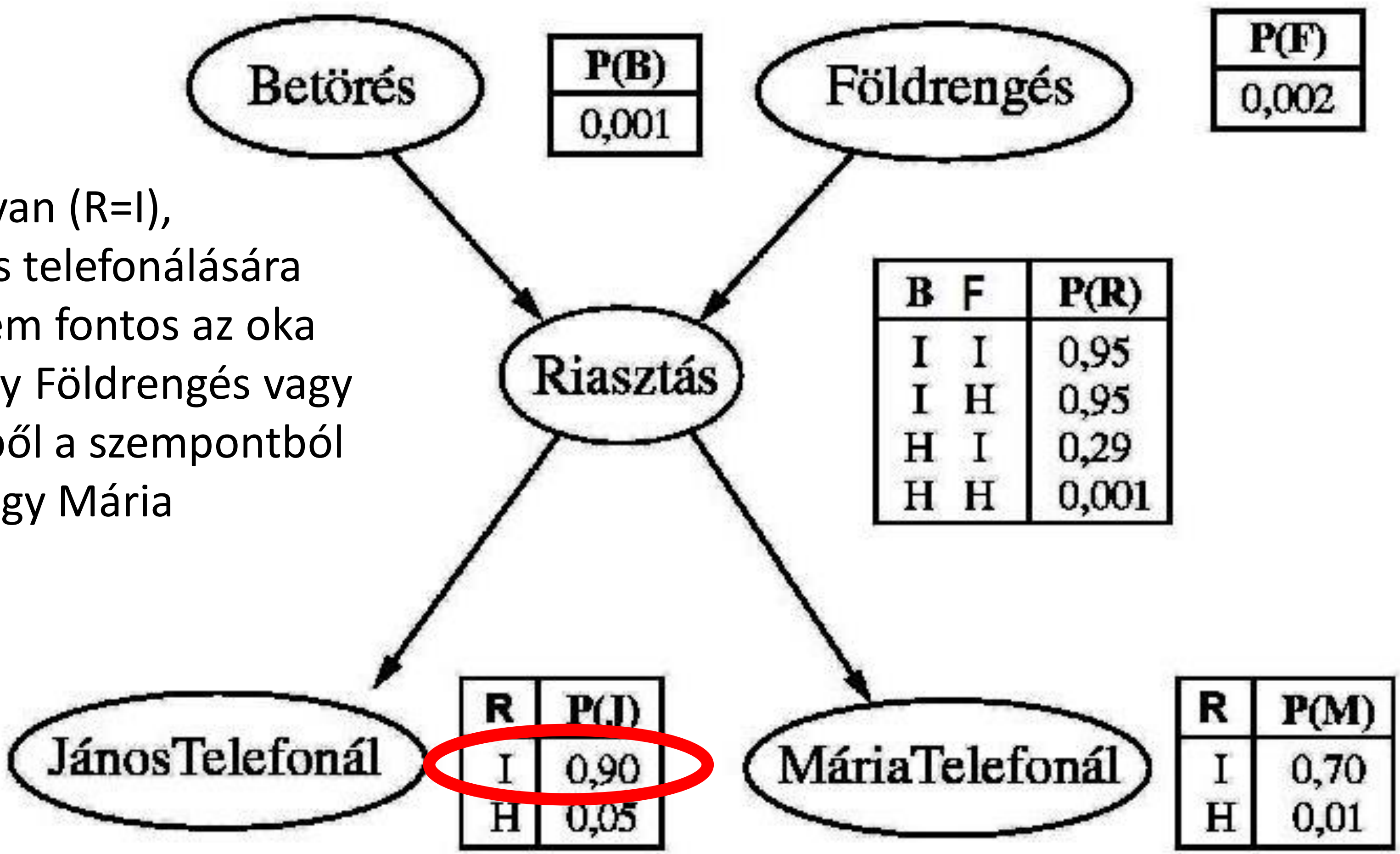
Mária csak 70% eséllyel.

Viszont néha más zajokat is riasztásnak gondolnak, és telefonálnak akkor is, amikor nincs riasztás. Ez Jánosnál 5% eséllyel fordul elő, Máriánál 1% eséllyel.

# Bayes (valószínűségi) hálók

Az együttes eloszlást $P(x_1, x_2, \ldots, x_N)$ -t adjuk meg, de kihasználjuk a változók közt fennálló feltételes függetlenségeket!

Ha pl. Riasztás van (R=I), akkor már János telefonálására vonatkozóan nem fontos az oka (Betörés és/vagy Földrengés vagy egyik sem)! Ebből a szempontból az se fontos, hogy Mária telefonál-e.




| P(B) |
|---|
| 0,001 |

| P(F) |
|---|
| 0,002 |

| B | F | P(R) |
|---|---|---|
| I | I | 0,95 |
| I | H | 0,95 |
| H | I | 0,29 |
| H | H | 0,001 |

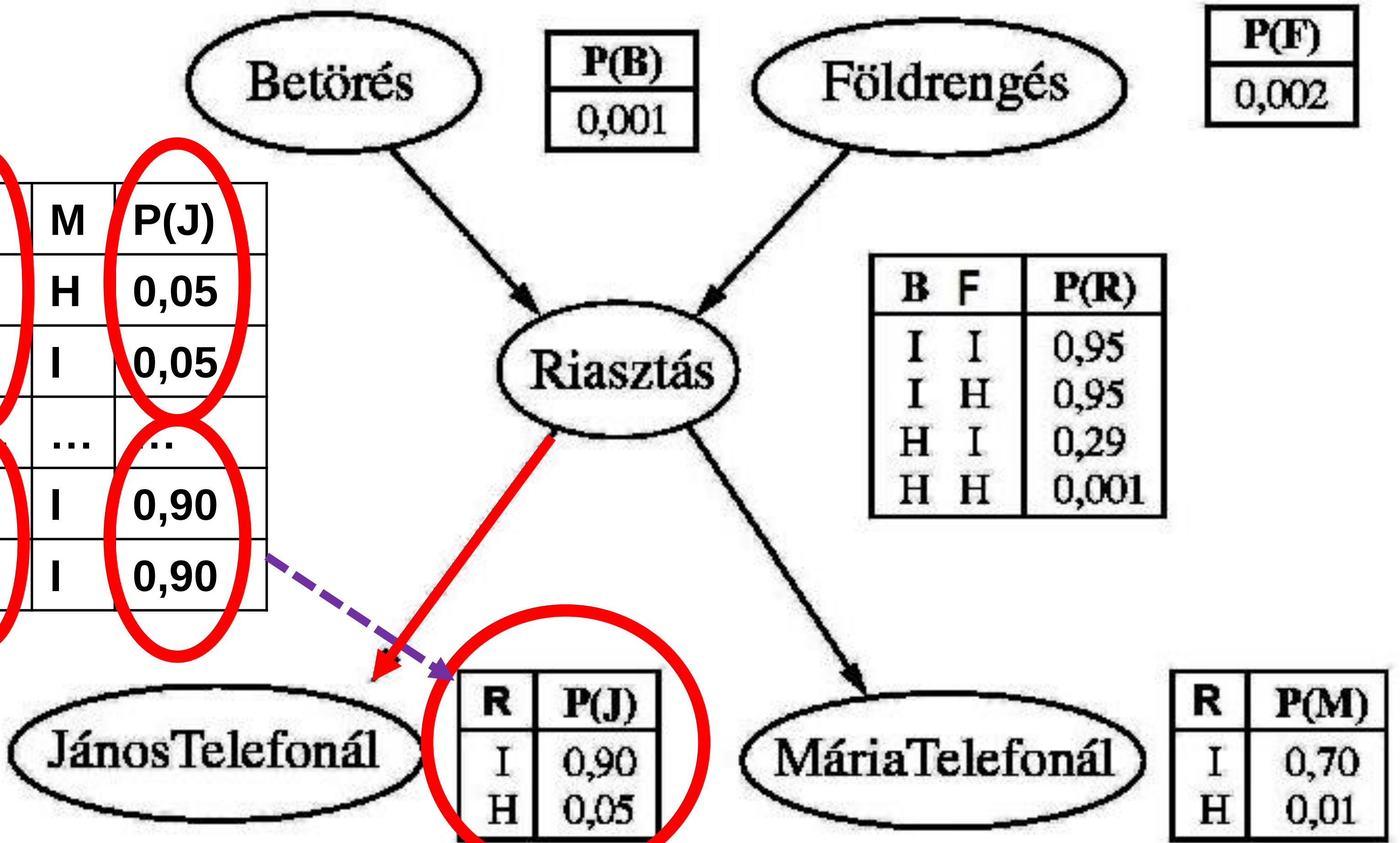
| R | P(J) |
|---|---|
| I | 0,90 |
| H | 0,05 |

| R | P(M) |
|---|---|
| I | 0,70 |
| H | 0,01 |

# Bayes (valószínűségi) hálók

Az együttes eloszlást $P(x_1, x_2, \ldots, x_N)$ -t adjuk meg, de kihasználjuk a változók közt fennálló feltételes függetlenségeket!

| B | F | R | M | P(J) |
|---|---|---|---|---|
| H | H | H | H | 0,05 |
| H | H | H | I | 0,05 |
| … | … | … | … | … |
| H | I | I | I | 0,90 |
| I | I | I | I | 0,90 |

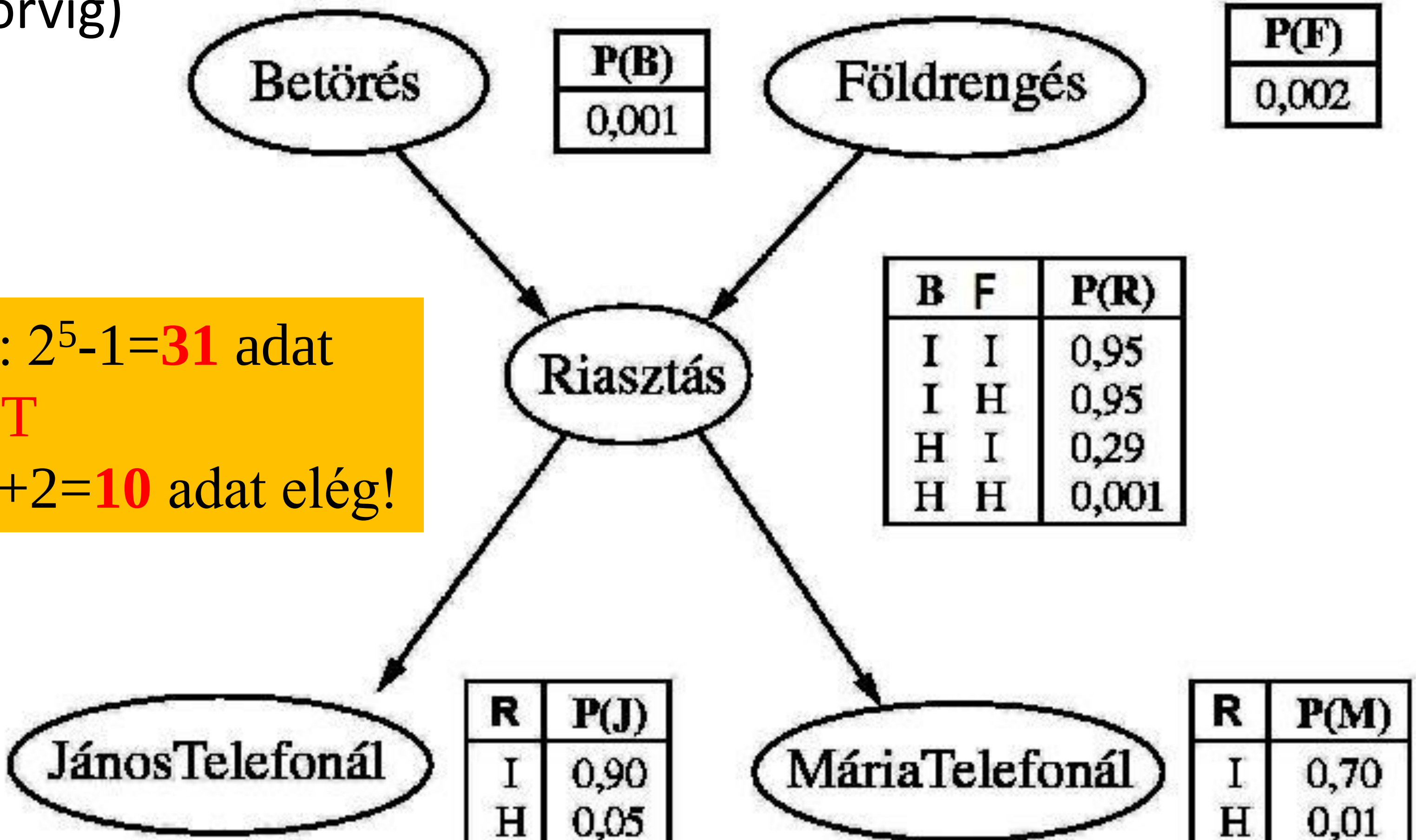


# Bayes (valószínűségi) hálók

Az együttes eloszlást $P(x_1, x_2, \ldots, x_N)$ -t adjuk meg, de kihasználjuk a változók közt fennálló feltételes függetlenségeket!

A jegyzet (Russell-Norvig) példája:

5 változó: $2^5$-1=**31** adat
HELYETT
1+1+4+2+2=**10** adat elég!




| P(B) |
|---|
| 0,001 |

| P(F) |
|---|
| 0,002 |

| B | F | P(R) |
|---|---|---|
| I | I | 0,95 |
| I | H | 0,95 |
| H | I | 0,29 |
| H | H | 0,001 |

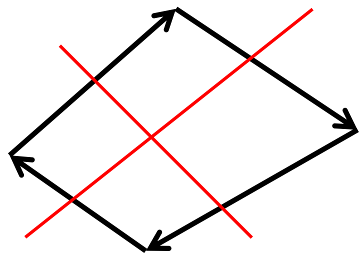
| R | P(J) |
|---|---|
| I | 0,90 |
| H | 0,05 |

| R | P(M) |
|---|---|
| I | 0,70 |
| H | 0,01 |

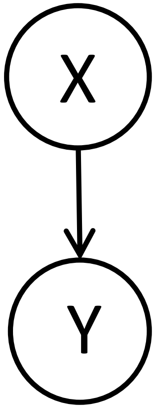
# Bayes (valószínűségi) hálók

Az együttes eloszlást $P(x_1, x_2, \ldots, x_N)$ -t adjuk meg, de kihasználjuk a változók közt fennálló feltételes függetlenségeket!

- A Bayes háló egy irányított gráf
- A háló csomópontjai a valószínűségi változók
- A csomópontokhoz valószínűségi információk vannak rendelve
- Bizonyos csomópontokat irányított élek kötnek össze. Az $X \rightarrow Y$ esetén X az Y-nak szülője
- A gráf nem tartalmaz irányított kört



- Minden csomóponthoz tartozik egy feltételes valószínűségeloszlás $P(X_k \mid Szülők(X_k))$




$P(X) = P(X=1)$ Az X bekövetkezésének valószínűsége

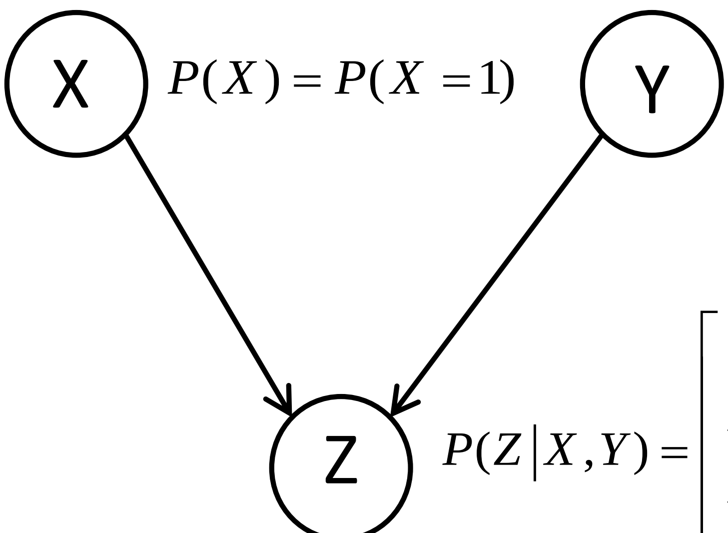
$$P(Y|X) = \begin{bmatrix} P(Y=1|X=1) \\ P(Y=1|X=0) \end{bmatrix}$$

pl. $P(Y=1|X=1)$ annak valószínűsége, hogy Y bekövetkezik akkor, ha X-ről tudjuk, hogy bekövetkezett.

Mivel: $P(Y=1|X=1) + P(Y=0|X=1) = 1$

tehát $\begin{bmatrix} P(Y=1|X=1) \\ P(Y=1|X=0) \end{bmatrix}$-ból $\begin{bmatrix} P(Y=0|X=1) \\ P(Y=0|X=0) \end{bmatrix} = \begin{bmatrix} 1-P(Y=1|X=1) \\ 1-P(Y=1|X=0) \end{bmatrix}$ megkapható

---




$P(X) = P(X=1)$ $P(Y) = P(Y=1)$

$$P(Z|X,Y) = \begin{bmatrix} P(Z=1|X=1,Y=1) \\ P(Z=1|X=1,Y=0) \\ P(Z=1|X=0,Y=1) \\ P(Z=1|X=0,Y=0) \end{bmatrix}$$

| X | Y | P(Z) |
|---|---|---|
| I | I | *P1* |
| I | H | *P2* |
| H | I | *P3* |
| H | H | *P4* |

# Következtetés valószínűségi hálókban

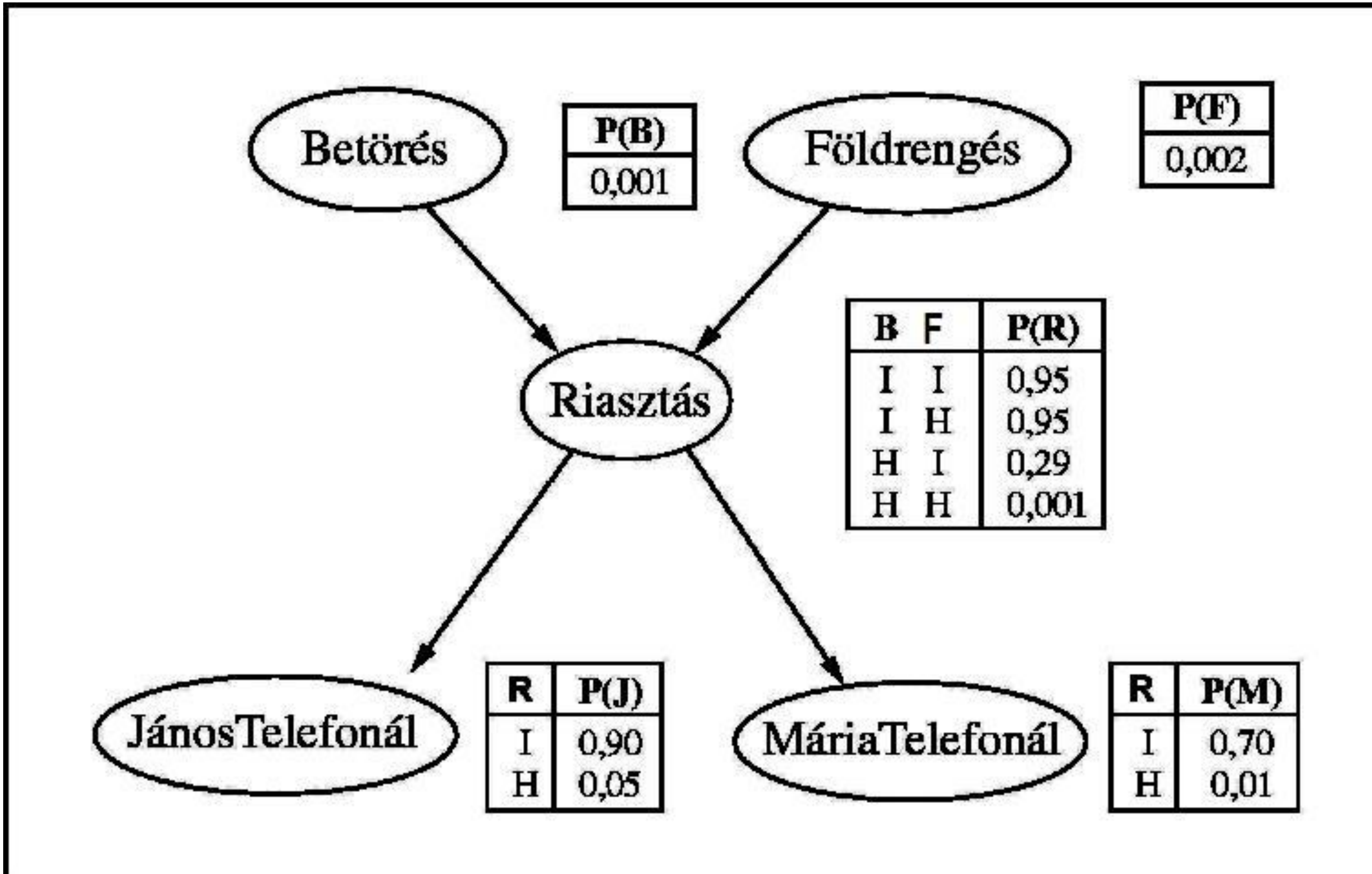
P(lekérdezés | bizonyíték) = ?

Valamire kíváncsiak vagyunk : lekérdezzük (egy *P*(*Változó*=Érték) valószínűségre kérdezünk rá).

A háló a rendelkezésére álló bizonyítékok alapján megadja, hogy az általunk kérdezett eseménynek mekkora a valószínűsége.

A jegyzet (Russell-Norvig) előbb már látott példáján mutatjuk be

A jegyzet (Russell-Norvig) példája:



Az ok$\rightarrow$okozati irányban viszonylag egyszerű a következtetés:
pl. ha adottak B és F valamilyen értékei, és ebből akarunk R-re, M-re vagy J-re következtetni. De hogyan következtethetünk „visszafele”?

Például kíváncsiak vagyunk a **P(J|M,¬B)** valószínűségre, tehát arra, hogy *mekkora valószínűséggel telefonál János, ha nem volt betörés, de Mária telefonál. A földrengésről és a riasztásról nem tudjuk, hogy bekövetkeztek-e….*

Bayes-tétel: $$P(J \mid M, \neg B) = \frac{P(J, M, \neg B)}{P(M, \neg B)}$$

**Összegzés rejtett (nem ismert értékű) változókra:**

$$\boxed{P(J, M, \neg B) = \sum_{r,f} P(J, M, \neg B, r, f)}$$

$$\boxed{P(M, \neg B) = \sum_{j,r,f} P(j, M, \neg B, r, f)}$$

Tehát példaként kíváncsiak vagyunk a **P(J|M,¬B)** valószínűségre….

Bayes-tétel:

$$P(J \mid M, \neg B) = \frac{P(J, M, \neg B)}{P(M, \neg B)}$$

Összegzés rejtett változókra:

$$P(J, M, \neg B) = \sum_{r, f} P(J, M, \neg B, r, f) =$$

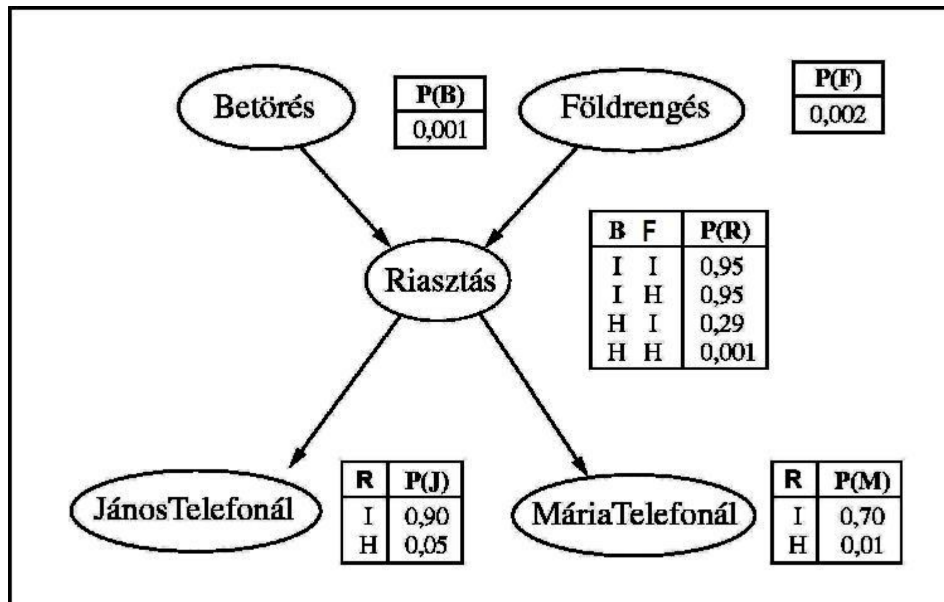
$$= P(J, M, \neg B, R, F) + P(J, M, \neg B, \neg R, F) +$$

$$+ P(J, M, \neg B, R, \neg F) + P(J, M, \neg B, \neg R, \neg F)$$

ezen belül az egyes tagok, pl.

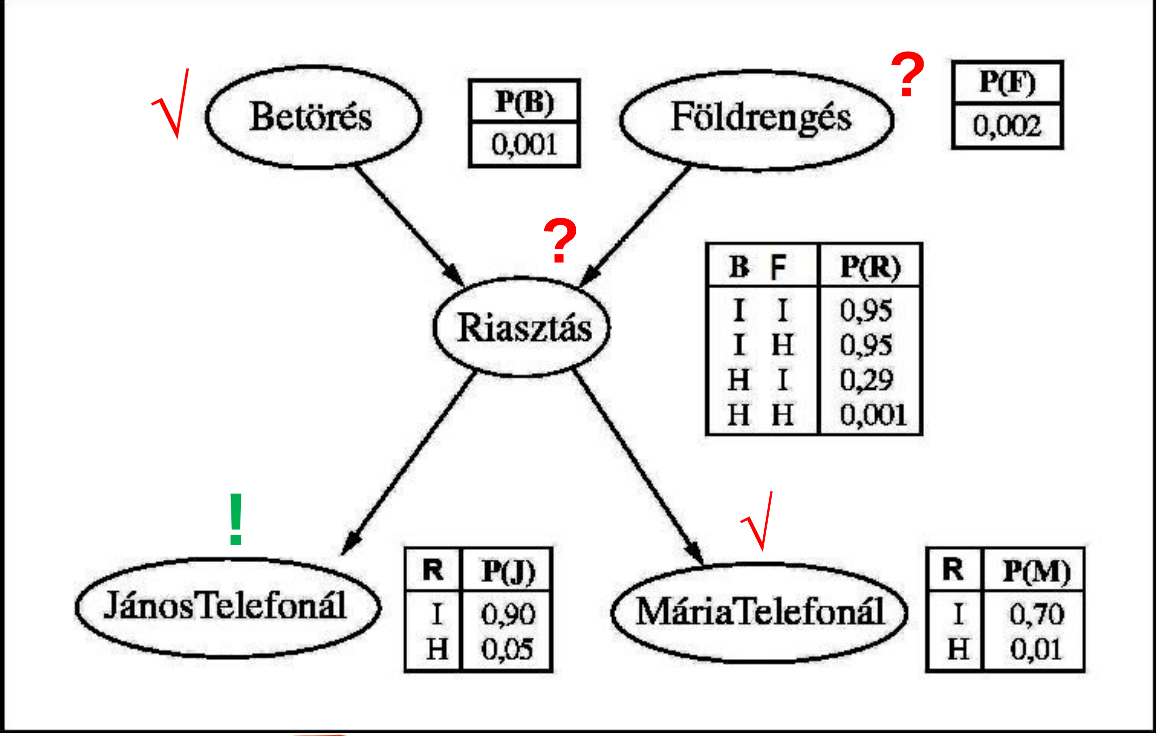
$$P(J, M, \neg B, R, F)$$

(a többi is erre a mintára megy)




$$P(J, M, \neg B, R, F) =$$

$$= P(J \mid R) \cdot P(M \mid R) \cdot P(R \mid \neg B, F) \cdot P(\neg B) \cdot P(F)$$




| P(B) |
|---|
| 0,001 |

| P(F) |
|---|
| 0,002 |

| B | F | P(R) |
|---|---|---|
| I | I | 0,95 |
| I | H | 0,95 |
| H | I | 0,29 |
| H | H | 0,001 |

| R | P(J) |
|---|---|
| I | 0,90 |
| H | 0,05 |

| R | P(M) |
|---|---|
| I | 0,70 |
| H | 0,01 |

## Összegzés rejtett változókra

Pl. $P(J|M,\neg B) = ?$

$= P(J,M,\neg B)/P(M,\neg B)$

$P(J,M,\neg B) =$

$\Sigma_{rf}\, P(J,M,\neg B, r, f)$

$P(M,\neg B) =$

$\Sigma_{rfj}\, P(M,\neg B, r ,f, j)$

$P(J\ M\neg B\ \ R\ \ F) = P(J|R)\ P(M|R)\ P(R|\neg BF)\ P(\neg B)\ P(F) = \ldots.$

$P(J\ M\neg B\ \neg R\ \ F) = P(J|\neg R)\ P(M|\neg R)\ P(\neg R|\neg BF)\ P(\neg B)\ P(F) = \ldots.$

$P(J\ M\neg B\ \ R\ \neg F) = P(J|R)\ P(M|R)\ P(R|\neg B\neg F)\ P(\neg B)\ P(\neg F) = \ldots.$

$P(J\ M\neg B\ \neg R\ \neg F) = P(J|\neg R)\ P(M|\neg R)\ P(\neg R|\neg B\neg F)\ P(\neg B)\ P(\neg F) = \ldots.$

$P(M\neg B\ \ J\ \ R\ \ F) = P(J|R)\ P(M|R)\ P(R|\neg BF)\ P(\neg B)\ P(F) = \ldots.$

$P(M\neg B\ \ J\ \ R\neg F) = P(J|R)\ P(M|R)\ P(R|\neg B\neg F)\ P(\neg B)\ P(\neg F) = \ldots.$

$P(M\neg B\ \ J\neg R\ \ F) = P(J|\neg R)\ P(M|\neg R)\ P(\neg R|\neg BF)\ P(\neg B)\ P(F) = \ldots..$

$P(M\neg B\ \ J\neg R\neg F) = P(J|\neg R)\ P(M|\neg R)\ P(\neg R|\neg B\neg F)\ P(\neg B)\ P(\neg F) = \ldots.$

$P(M\neg B\ \neg J\ \ R\ \ F) = P(\neg J|R)\ P(M|R)\ P(R|\neg BF)\ P(\neg B)\ P(F) = \ldots..$

$P(M\neg B\ \neg J\ \ R\neg F) = P(\neg J|R)\ P(M|R)\ P(R|\neg B\neg F)\ P(\neg B)\ P(\neg F) = \ldots..$

$P(M\neg B\ \neg J\neg R\ \ F) = P(\neg J|\neg R)\ P(M|\neg R)\ P(\neg R|\neg BF)\ P(\neg B)\ P(F) = \ldots..$

$P(M\neg B\ \neg J\neg R\neg F) = P(\neg J|\neg R)\ P(M|\neg R)\ P(\neg R|\neg B\neg F)\ P(\neg B)\ P(\neg F) = \ldots.$

# **Következtetés felsorolással** (balról jobbra)

Legyen most a példánk: ha János is, Mária is telefonált, akkor mi a Betörés valószínűsége?

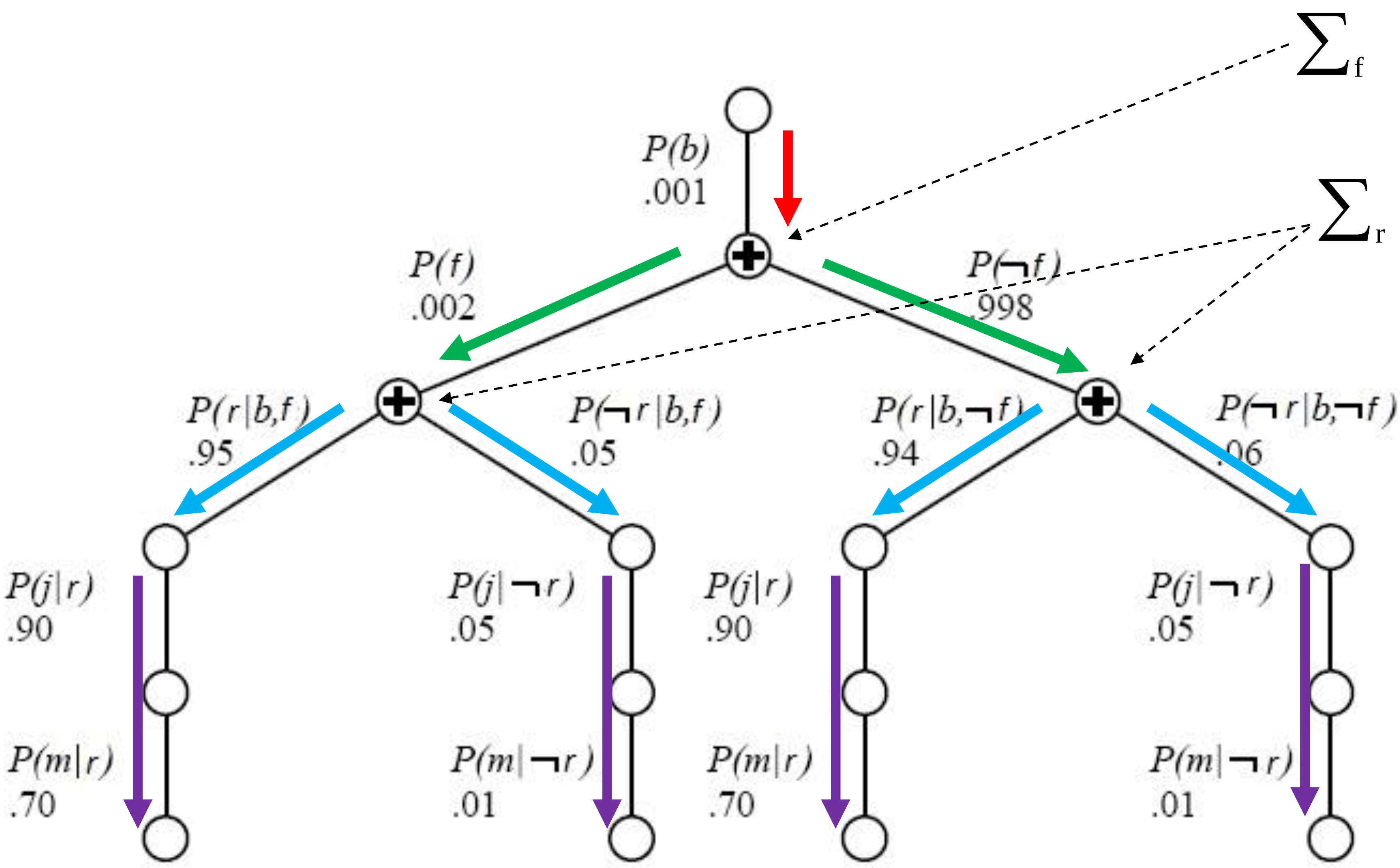
$$P(\mathrm{B} \mid \mathrm{J}, \mathrm{M}) = P(b{=}\mathrm{Igaz} \mid j{=}\mathrm{Igaz},\ m{=}\mathrm{Igaz}) = ?$$

$$P(\mathrm{B} \mid \mathrm{J}, \mathrm{M}) = \Sigma_f\, \Sigma_r\, \mathrm{P}(\mathrm{B}, f, r, \mathrm{J}, \mathrm{M})$$

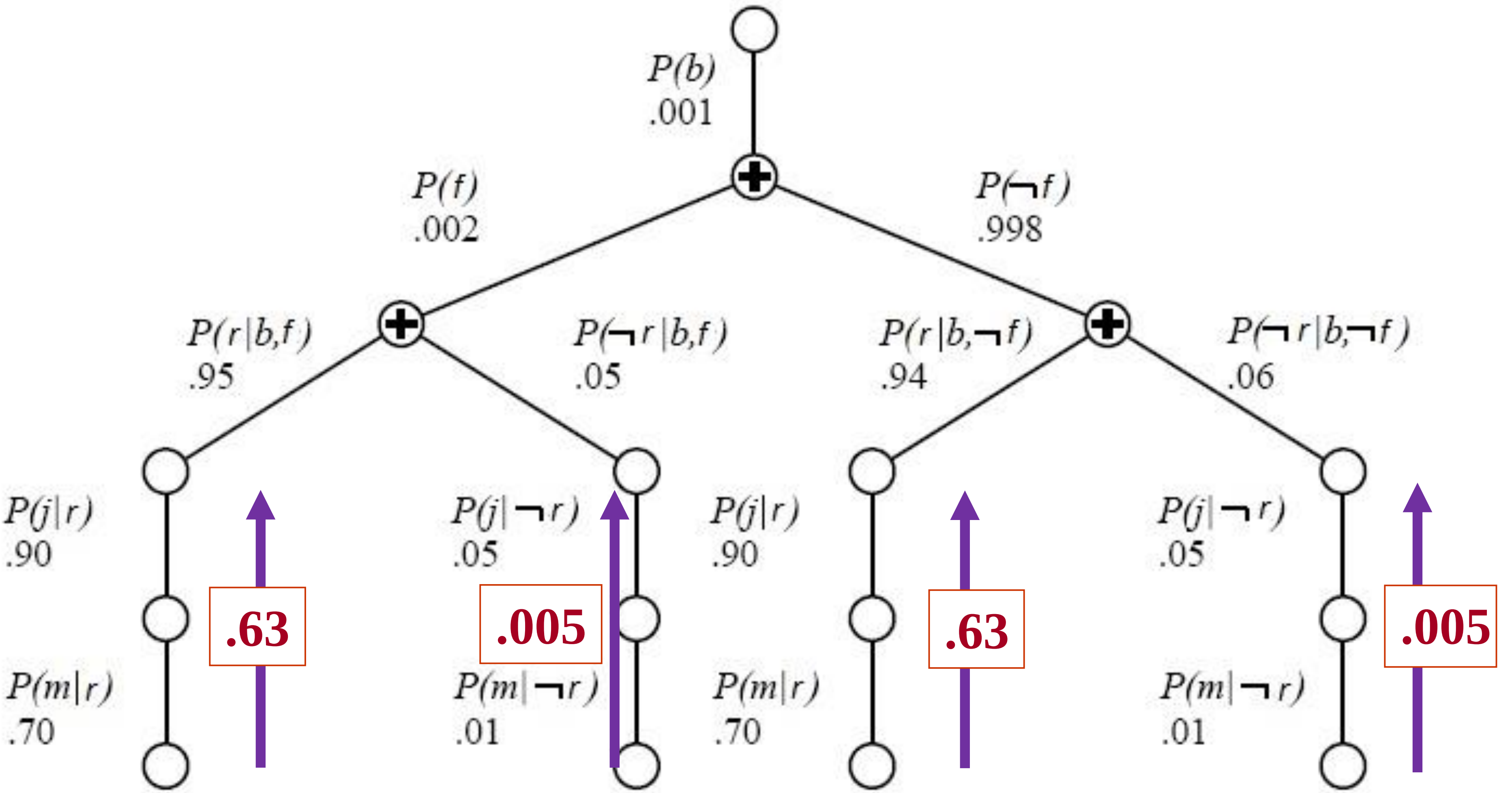
$$= \Sigma_f \quad \Sigma_r \quad \mathrm{P}(\mathrm{B})\ \mathrm{P}(f) \quad \mathrm{P}(r \mid \mathrm{B}, f)\ \mathrm{P}(\mathrm{J} \mid r)\ \mathrm{P}(\mathrm{M} \mid r)$$

$$= \mathrm{P}(\mathrm{B}) \quad \Sigma_f \quad \mathrm{P}(f) \quad \Sigma_r\, \mathrm{P}(r \mid \mathrm{B}\ f)\ \mathrm{P}(\mathrm{J} \mid r)\ \mathrm{P}(\mathrm{M} \mid r)$$

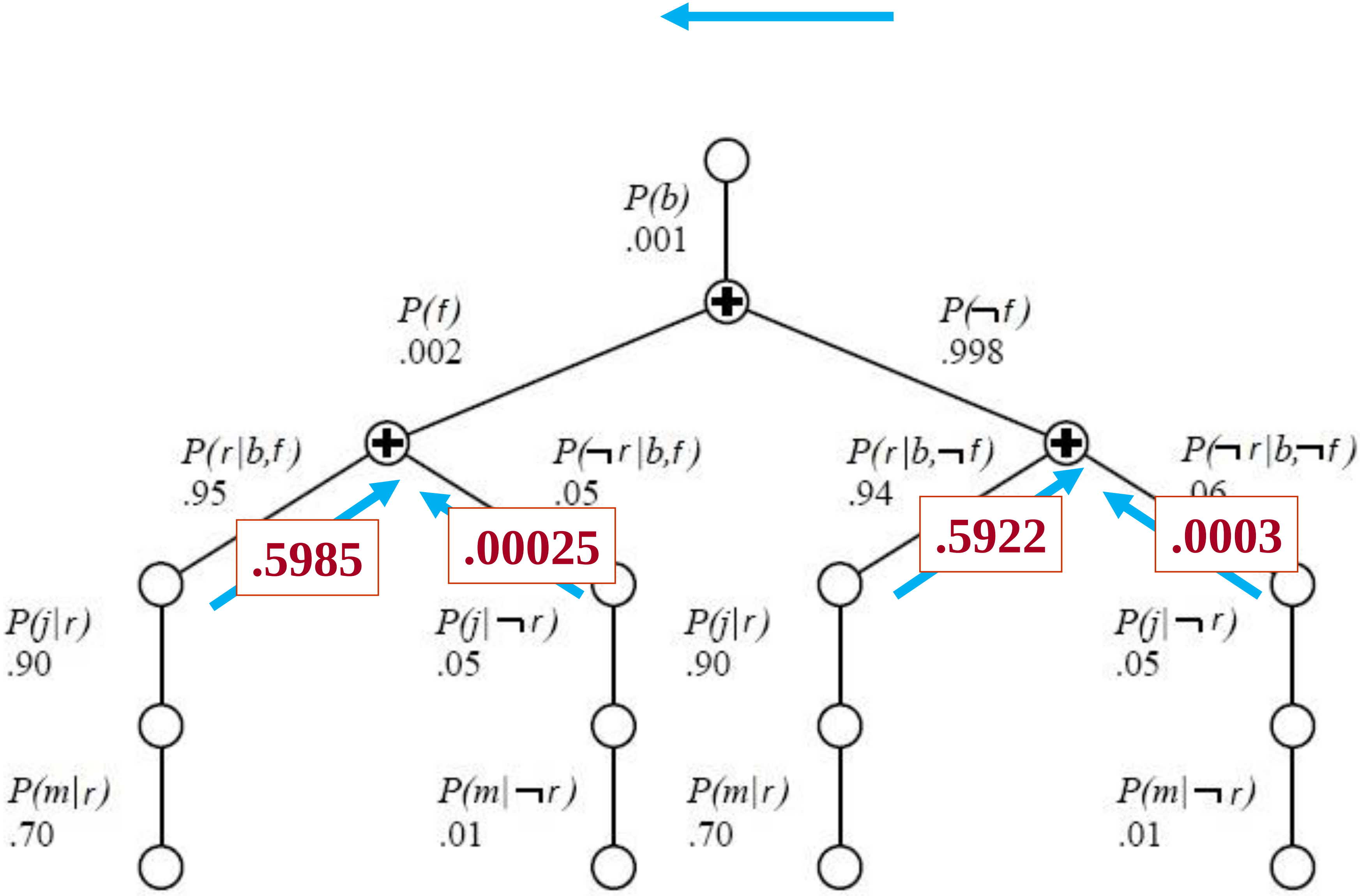
$$P(B \mid J, M) = P(B)\ \Sigma_f\ P(f)\ \Sigma_r\ P(r \mid B\ f)\ P(J \mid r)\ P(M \mid r)$$



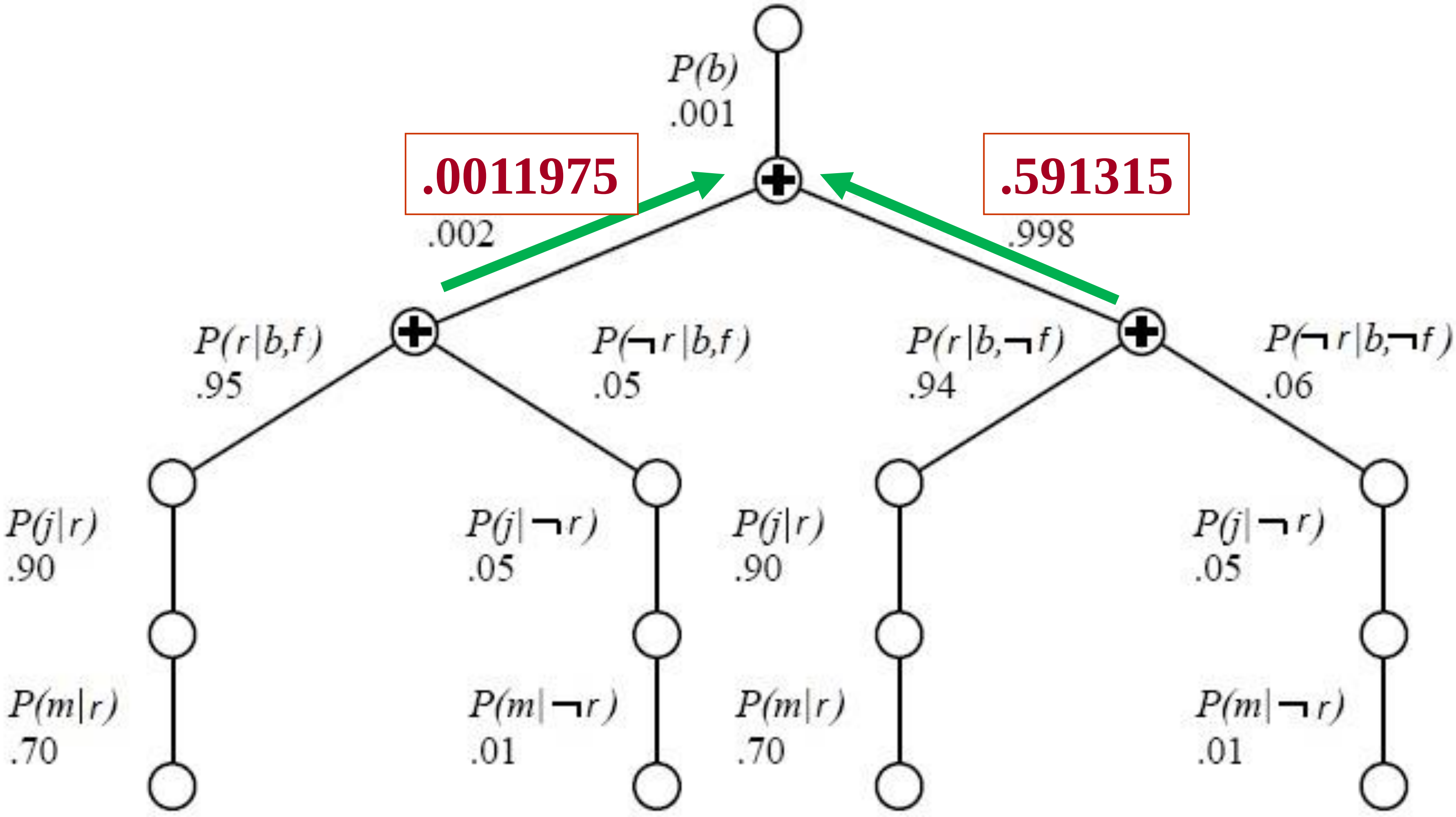
$$P(B \mid J\ M) = \alpha\, P(B)\, \Sigma_f\, P(f)\, \Sigma_r\, P(r \mid B\ f)\, P(J \mid r)\, P(M \mid r)$$



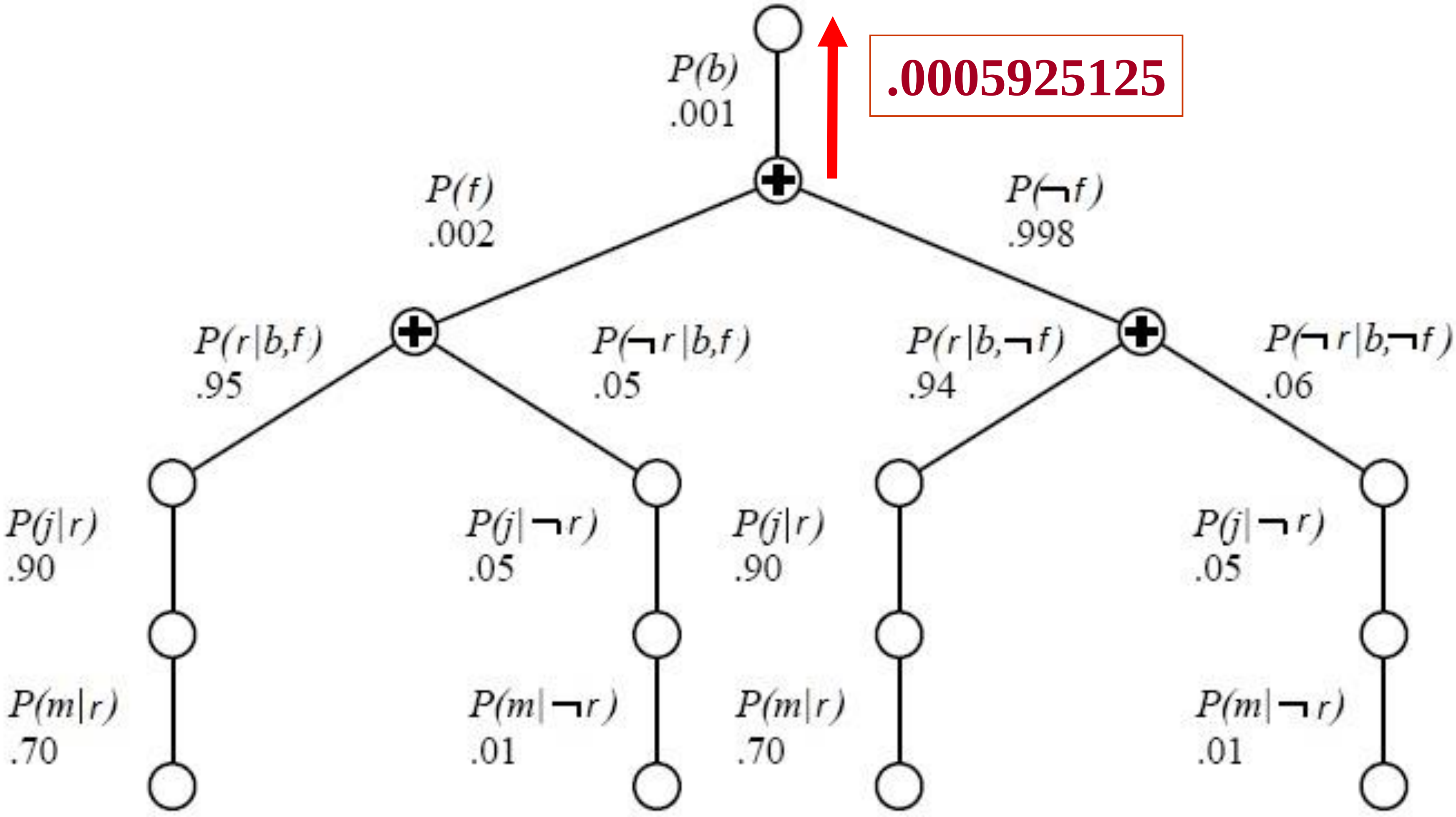
$$P(B \mid J\ M) = \alpha\ P(B)\ \Sigma_f\ P(f)\ \Sigma_r\ P(r \mid B\ f)\ P(J \mid r)\ P(M \mid r)$$



$$P(B \mid J\ M) = \alpha\ P(B)\ \Sigma_f\ P(f)\ \Sigma_r\ P(r \mid B\ f)\ P(J \mid r)\ P(M \mid r)$$



$$P(B \mid J\ M) = P(B)\ \Sigma_f\ P(f)\ \Sigma_r\ P(r \mid B\ f)\ P(J \mid r)\ P(M \mid r)$$

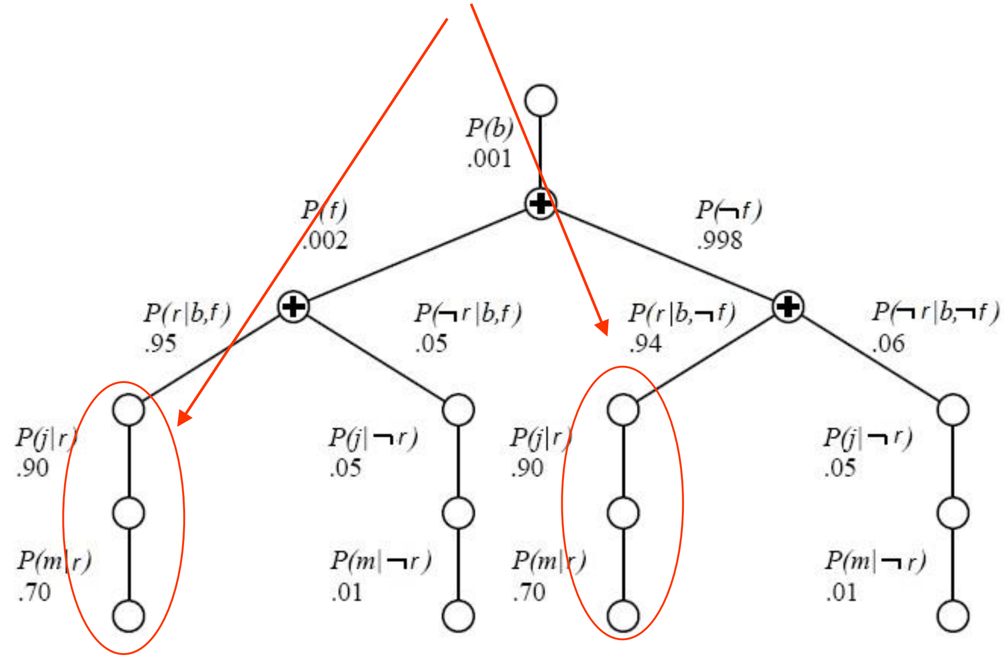


$$P(B \mid J,M) = P(B)\, \Sigma_f\, P(f)\, \Sigma_r\, P(r \mid B\ f)\, P(J \mid r)\, P(M \mid r)$$

**Kiértékelés**:

- most: jobbról balra
- köztes eredmények tárolása

Az előbbi megoldás problémája:
bizonyos tagok többszörös kiszámítása, pl.



# Irreleváns változók eliminálása

*Vigyázat! új feladat (kérdés) van ebben a példában:*

$$P(J \mid B) = \alpha\, P(B) \sum_f P(f) \sum_r P(r \mid B\, f)\, P(J \mid r) \overbrace{\sum_m P(m \mid r)}^{= 1}$$

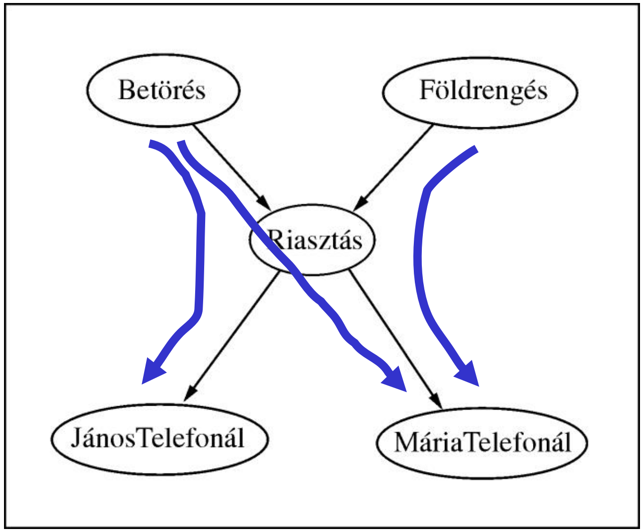
P(*JánosTelefonál*|*Betörés*) lekérdezés: *f*-re, *r*-re és *m*-re kell összegeznünk, de az eredményt nem változtatja meg a *MáriaTelefonál* eltávolítása a hálóból.

Általában, bármely **levél** csomópontot eltávolíthatunk, ami **nem célváltozó** vagy **nem bizonyíték** változó.

**Eltávolítás után lehetnek újabb levélcsomópontok, amelyek szintén irrelevánsak lehetnek**.

**Minden változó, ami nem őse a célváltozónak vagy egy bizonyíték változónak, irreleváns a lekérdezésre.** A változó elimináló algoritmus ezért az összes ilyen változót eltávolíthatja a lekérdezés kiértékelése előtt.

# Az egzakt következtetés komplexitása

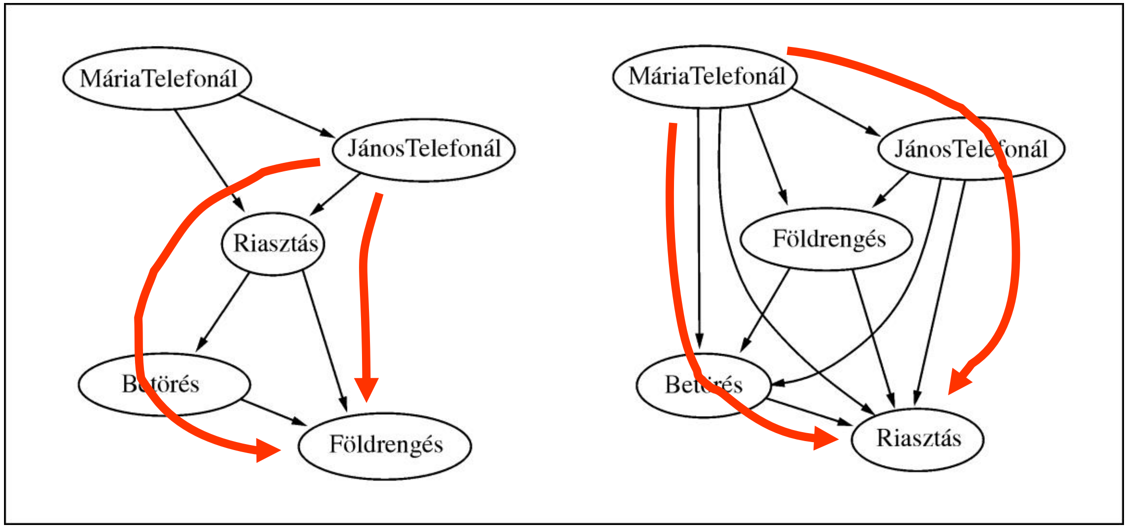



**egyszeresen összekötött,**

bármely két csp. között legfeljebb egyetlen egy út van

**többszörösen összekötött**

bármely két csomópont között több út van



## Algoritmusok lekérdezések megválaszolására

**A.) egyszeresen összekötött, fa gráf**

- létezik lineáris komplexitású algoritmus

**B.) többszörösen összekötött**

- nem létezik zárt alakú algoritmus
- vagy a komplexitás megugrik (lásd előbb), vagy csak közelítő módszerek maradnak

Sokszor *jobb egy* gyorsan kiértékelhető *közelítő módszer*, *mint egy* nagyon lassú, komplex pontos módszer! (Mérnökök vagyunk, nem matematikusok…)