Budapesti Műszaki és Gazdaságtudományi Egyetem
Villamosmérnöki és Informatikai Kar
Mesterséges Intelligencia és Rendszertervezés Tanszék

Mesterséges intelligencia előadássorozat

Az előadás diái az AIMA könyvre épülve (http://aima.cs.berkeley.edu) készültek a University of California, Berkeley mesterséges intelligencia kurzusának anyagainak felhasználásával (http://ai.berkeley.edu).

These slides are based on the AIMA book (http://aima.cs.berkeley.edu) and were adapted from the AI course material of University of California, Berkeley (http://ai.berkeley.edu).

MŰEGYETEM 1782

Budapesti Műszaki és Gazdaságtudományi Egyetem
Villamosmérnöki és Informatikai Kar
Mesterséges Intelligencia és Rendszertervezés Tanszék

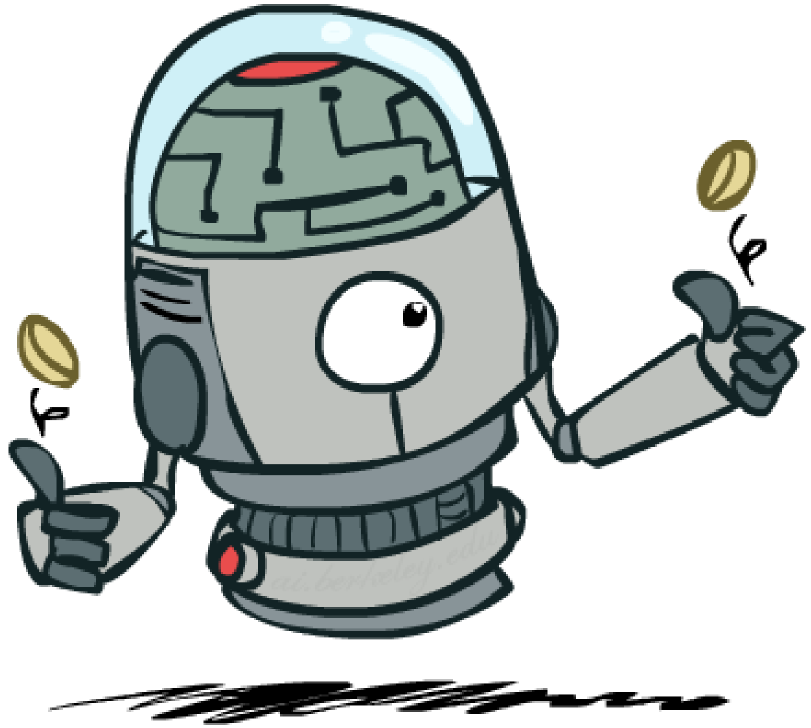
Valószínűségi hálók

Feltételes függetlenségek,
kompakt reprezentáció, következtetés

Előadó: Dr. Hullám Gábor
Előadás anyaga: Dr. Antal Péter

MŰEGYETEM 1782

Függetlenség



Függetlenség

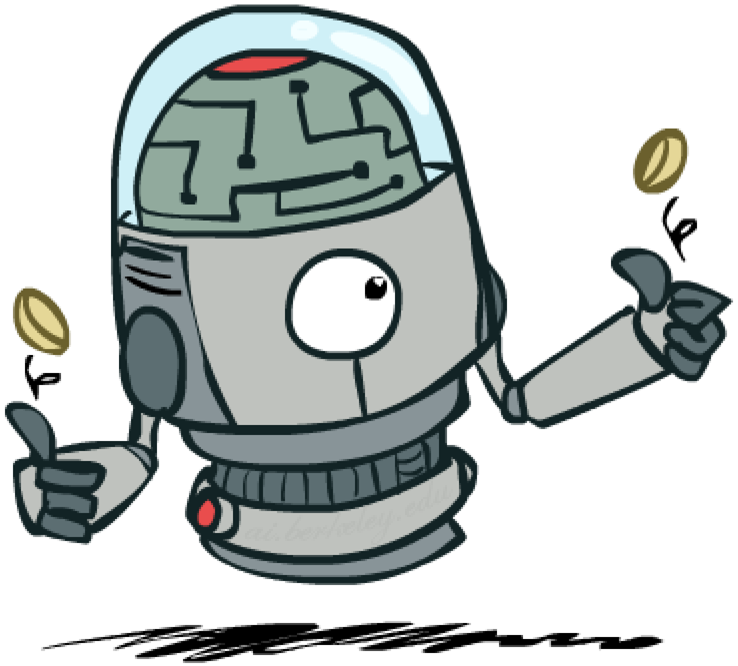
- Két valószínűségi változó független, ha

$$\forall x, y : P(x, y) = P(x)P(y)$$

  - azaz az együttes eloszlás két **egyszerűbb** eloszlás szorzata.
- Alternatív feltétel:

$$\forall x, y : P(x|y) = P(x)$$

  - azaz a feltételes eloszlás helyett elég egy **egyszerűbb** eloszlás
- **A függetlenség teszi lehetővé az egyszerűsítést:**
  - komplex vélekedések modulárisan felbonthatóak
    - X,Y együttes eloszlása helyett elég az X és Y marginális eloszlása!
  - komplex következtetésekben evidenciák figyelmen kívül hagyhatóak
    - X vonatkozásában Y ismerete szükségtelen!



Független - Nem független

$P(H,I)$	I=napsütés	I=eső
H=meleg	0.3	0.1
H=hideg	0.4	0.2

I kiösszegzése

$P(H)$

H	P
meleg	0.4
hideg	0.6

H kiösszegzése

$P(I)$

I	P
napsütés	0.7
eső	0.3

Függetlenség feltevésével:
$P'(H,I)=P(H)*P(I)$

$P'(H,I)$	I=napsütés	I=eső
H=meleg	0.28	0.12
H=hideg	0.42	0.18

Függetlenség: releváns - irreleváns

A valószínűség szubjektív értelmezése esetén a függetlenség információs irrelevanciát jelent: Y ismerete nem változtatja meg X lehetséges értékeiről a vélekedést.

- (Szimmetrikus)
- Mi releváns, mi nem?
- El tudjuk a valószínűségek használata nélkül is dönteni?

A függetlenség, mint strukturális tulajdonság

- A függetlenségek az eloszlásokat kvalitatívan jellemzik
  - Adott függetlenségi reláció(k) eloszlások sokaságában teljesülnek
    - Valószínűségi feltevés:
      - 0<=x,y,z és x+y+z<=1

P(H,I)	I=napsütés	I=eső
H=meleg	x	y
H=hideg	z	1-(x+y+z)

0<=u,v<=1

P(I)	P
napsütés	v
eső	1-v

Függetlenség:
*P'(H,I)=P(H)*P(I)*

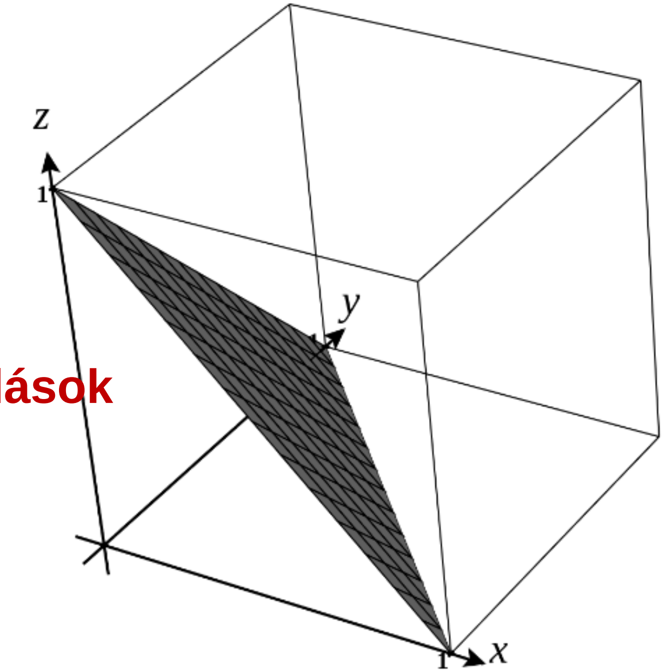
P(H)	P
meleg	u
hideg	1-u

P'(H,I)	I=napsütés	I=eső
H=meleg	u*v	u*(1-v)
H=hideg	(1-u)*v	(1-u)*(1-v)

A függetlenség ritka

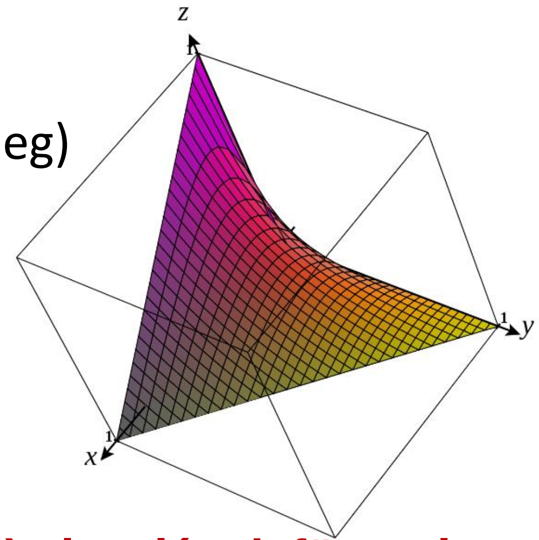
$P(H,I)$	I=napsütés	I=eső
H=meleg	x	y
H=hideg	z	1-(x+y+z)

- A függetlenség egy egyenletrendszer megoldása
  - Valószínűségi feltevés:
    - 0<=x,y,z és x+y+z<=1
  - Függetlenségi feltevés:
    - P(I=napsütés)=P(I=napsütés|H=meleg)
    - azaz: x+z = x/(x+y)



P(H,I) eloszlások

Okamoto 1973
Meek 1995



P(H,I) eloszlások független H,I-vel

A függetlenség ritka és sokat ér

A szabad paraméterek száma n bináris valószínűségi változó esetén

Általános eset

2^n

X_1	X_2	…	X_n	P
0	0	…	0	..
0	0		1	
…				
1	1		1	…

2ⁿ-1 darab szabad paraméter

Teljes függetlenség esetén

X_1	P
0	p_1
1	$1-p_1$

…

X_n	P
0	p_n
1	$1-p_n$

n darab szabad paraméter

Költség: szakértői idő vagy adatmennyiség becsléshez

Alapeseti eloszlás

- A függetlenség egy univerzálisan kantált állítás

$$\forall x, y : P(x|y) = P(x)$$

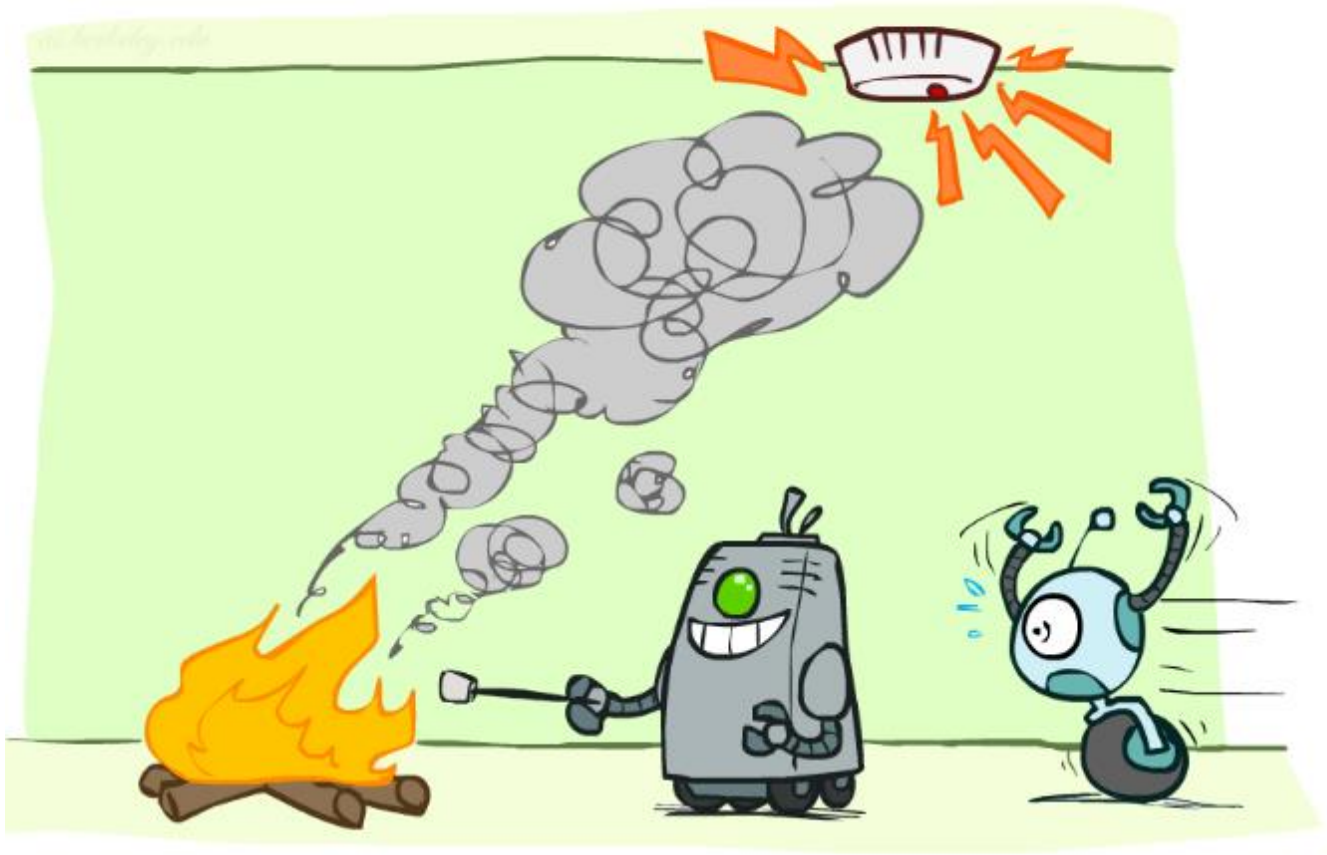
  - Lehetséges, hogy csak bizonyos *y* értékekre áll fenn szabályszerűség, pl

$$\exists y, y' : P(X|Y = y) = P(X|Y = y')$$

$$\underbrace{\hphantom{xxxxxxxxxxxxxxxxxxxxxxxxxxxxxx}}^{\text{Bevezethető egy alapeseti feltételes eloszlás}}$$

	I=napsütés	I=eső	I=tornádó	I=szélvihar	..
H=meleg	0.49	0.1	0.001	0.002	..
H=hideg	0.1	0.2	0.01	0.02	...

Feltételes függetlenség



Feltételes függetlenség

- *A feltételes függetlenség a* függetlenség gyakoribb formája.
- X feltételesen független Y-tól adott/ismert Z esetén:

$$\forall x, y, z : P(x, y|z) = P(x|z)P(y|z)$$

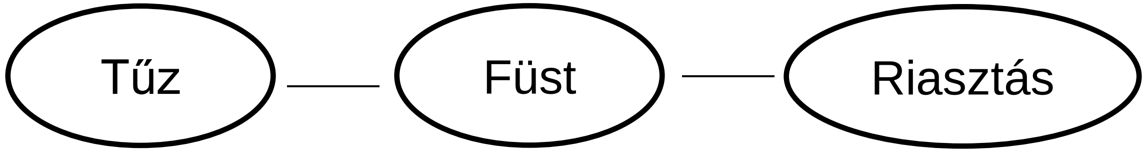
avagy,

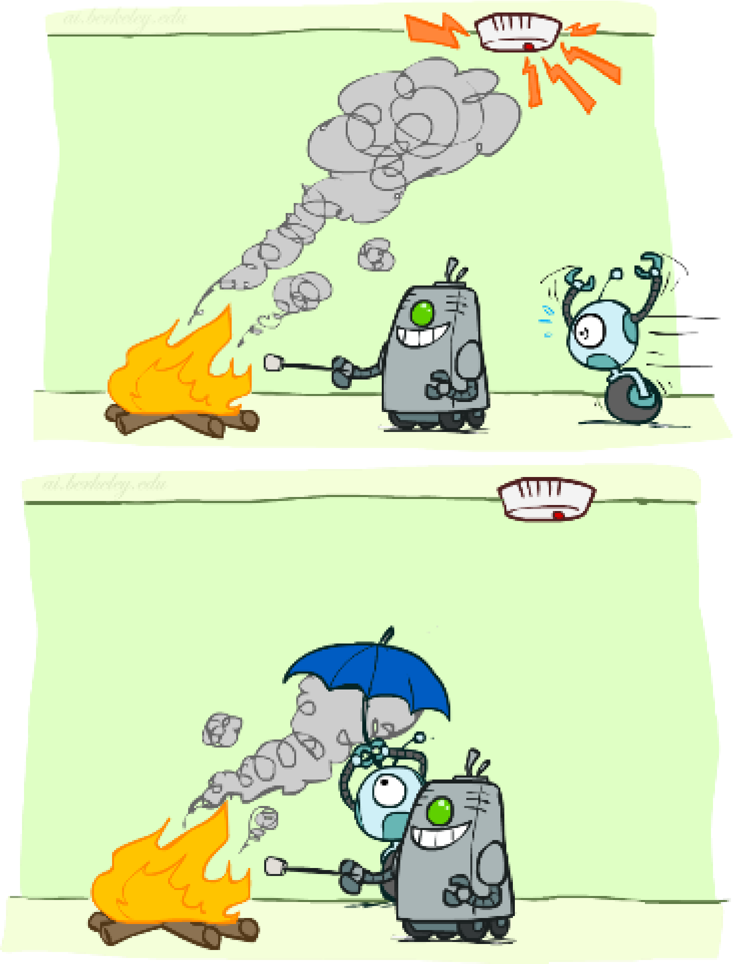
$$\forall x, y, z : P(x|z, y) = P(x|z)$$

- Jelölése:
  - $I_P(X;Y|Z)$
  - függésé: $D_P(X;Y|Z) = \neg I_P(X;Y|Z)$

Feltételes függetlenség

- Függetlenségek?
  - Tűz (T/t): nincs, kis, nagy
  - Füst (F/f): nincs, kis, nagy
  - Riasztás (R/r): nincs, van
  - Együttes eloszlás: P(T,F,R)
  - $P(T,R|F) \overset{?}{=} P(T|F)\,P(R|F)$

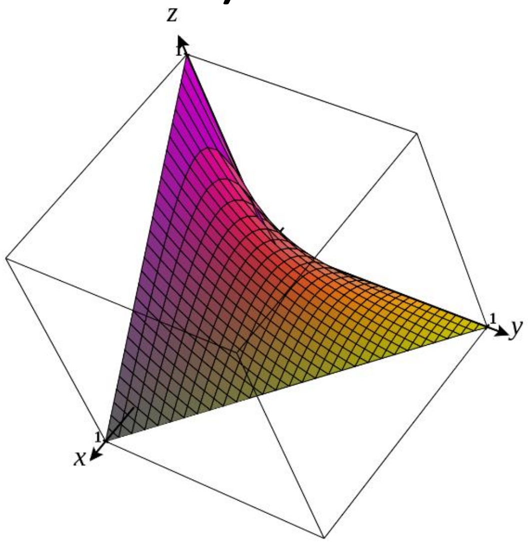




- Ha teljesül:
  - $P(T|F) = P(T|F,R)$
  - $P(R|F) = P(R|F,T)$
- **Kevesebb paraméter!**
  - ahogyan függetlenségnél is
- **Valószínűségi következtetés számok nélkül!!**
  - mi releváns, mi nem

A feltételes függetlenségek (is) ritkák

- A feltételes függetlenségek (is) egy egyenletrendszer megoldásai
  - Valószínűségszámításbeli általános feltevések
  - Feltételes függetlenségi feltevések
- ==> Adott feltételes függetlenséget teljesítő együttes eloszlások alacsonyabb dimenziójú felületeken helyezkednek el (szabad paramétereik száma kevesebb).

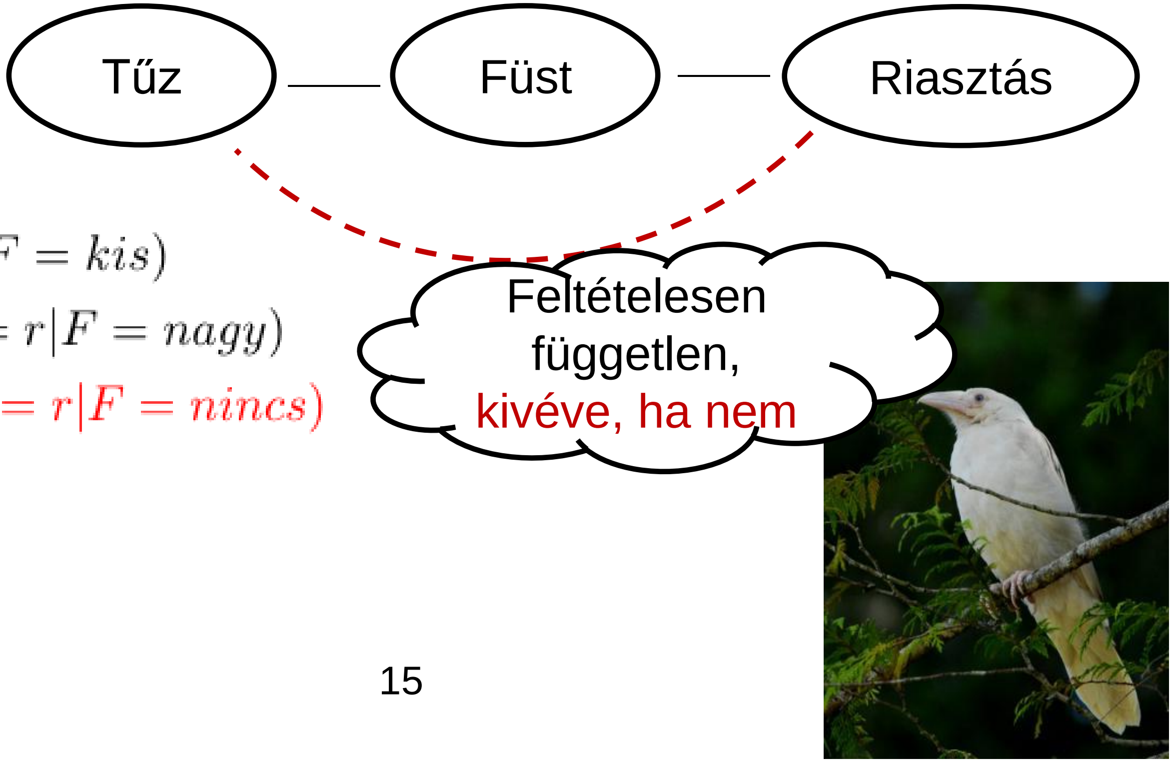


Kontextuális függetlenségek

- A feltételes függetlenség (is) egy univerzálisan kvantált állítás

$$\forall t, f, r : P(T=t, R=r|F=f) = P(T=t|F=f)P(R=r|F=f)$$

  - Lehetséges, hogy csak bizonyos f_0 kontextus esetén áll fenn a függetlenség
    - például, ha nincs füst, akkor a tűz maga is aktiválhatja a riasztót:



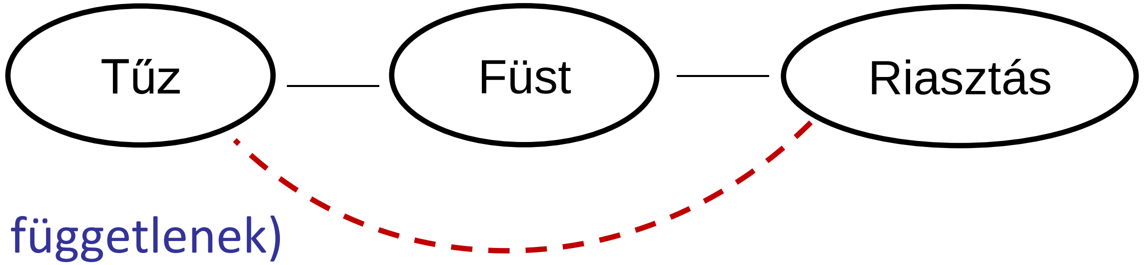
$$\forall t, r : P(T=t, R=r|F=kis) = P(T=t|F=kis)P(R=r|F=kis)$$

$$\forall t, r : P(T=t, R=r|F=nagy) = P(T=t|F=nagy)P(R=r|F=nagy)$$

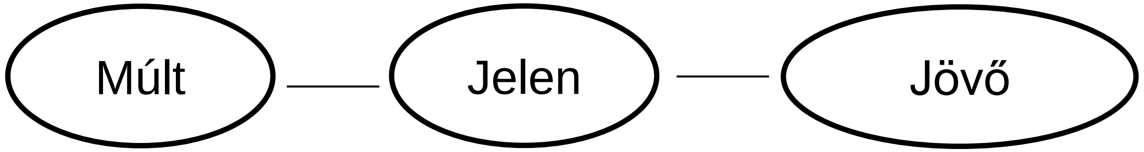
$$\forall t, r : P(T=t, R=r|F=nincs) \neq P(T=t|F=nincs)P(R=r|F=nincs)$$

Az intranzitív függés esete

- Lehet-e függések láncolata független?
- Lehet-e releváns információk láncolatából álló magyarázat irreleváns?
- Lehet-e oksági mechanizmusok láncolata teljesen hatástalan?
  - Azaz a példában
    - Tűz-Füst NEM független
    - Füst-Riasztás NEM független
    - Lehet-e a Tűz és Riasztás független?
      - (azt tudjuk, hogy Füst ismeretében feltételesen függetlenek)

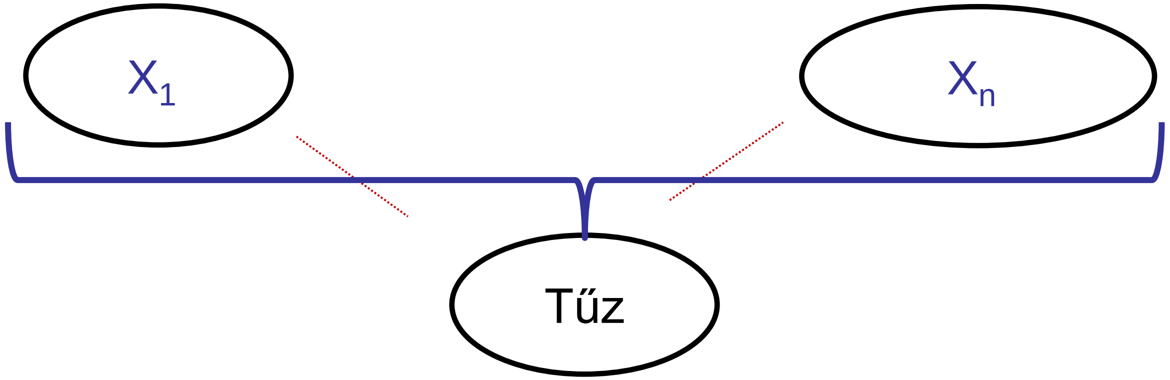


LEHETSÉGES!



A szinergia esete

- Létrejöhet-e függetlenségekből függés?
- Lehet-e irreleváns információk együttese releváns?
- Lehetnek-e okok önmagukban hatástalanok?
  - Tételezzük fel, hogy Tűz üt ki, ha egy tűzhely kapcsolóján páratlan számút fordítanak, azaz
    - jelölje X_i bináris változó egyenletes eloszlással, hogy az i. használó fordít-e
    - és legyen $Tűz=XOR(X_1, X_2,\ldots X_n)$.
    - Ekkor Tűz és bármely X_i független,
    - de Tűz (determinisztikusan!) függ X_i-k összességétől

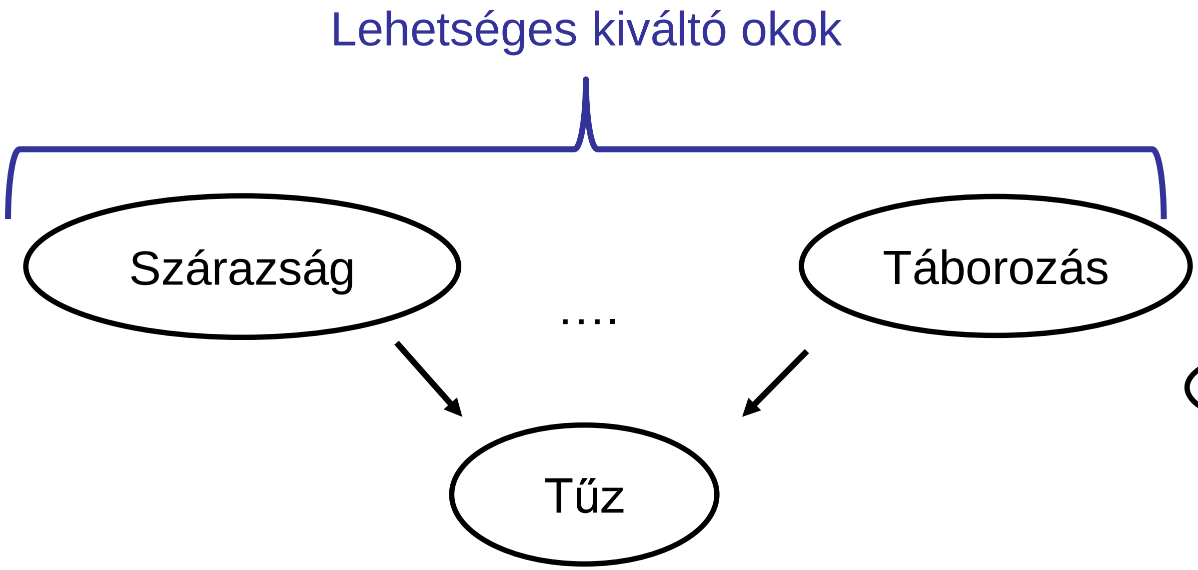


És ez is ritka...

Az idő nyilának története

- Lehetnek-e releváns információmorzsák egymásra nézve irrelevánsak?
- Lehetnek-e okok egymástól teljesen függetlenek?

TERMÉSZETESEN!



Strukturálisan ritka:
unikális

A közös következménnyel, mint feltétellel, már nem függetlenek!
("kimagyarázás" jelensége)

Függőségi modellek

- Egy bizonytalan probléma strukturális jellegzetességeit függetlenségekkel leírhatjuk
- Egy P együttes valószínűségi eloszlás függőségi modellje a P-ben érvényes feltételes függetlenségek halmaza:
  - $M_P=\{I_{P,1}(X_1;Y_1|Z_1),\ldots, I_{P,K}(X_K;Y_K|Z_K)\}$

P(X,Y,Z) egy (elsőrendű) Markov-lánc, ha
$M_P=\{D(X;Y), D(Y;Z), I(X;Z|Y),$ **D(X;Z)**$\}$

$X \rightarrow Y \rightarrow Z$

Múlt → Jelen → Jövő

Általában, de ritkán **I(X;Z)**

Függőségi modellek tulajdonságai

- Egy adott eloszlásban érvényes feltételes függetlenségek egy rendszert alkotnak logikai szabályszerűségekkel
  - Ha Z ismeretében X érdektelen információ Y-ra nézve, akkor X bármely X' része és aspektusa érdektelen:
    - Ha I(X;Y|Z), akkor I(X';Y|Z) és I(f(X);Y|Z)
- Lehet-e egy helyes és teljes logikai kalkulus függetlenségekre?
- Lehet-e gráfokat használni függetlenségek egzakt reprezentálásra?

Általában NEM
de majdnem mindig lehet

Ritkán
nem lehet

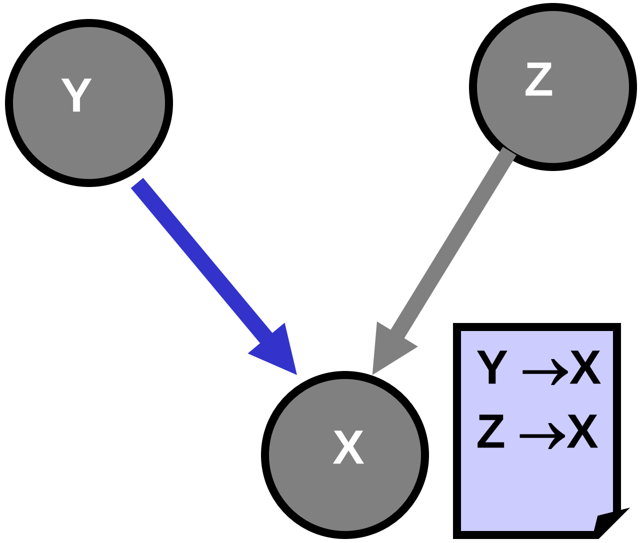
Összefoglalás

- A függetlenség a valószínűségi modellezési fő eszköze
  - egyszerűbb (kevesebb paraméterű) modellek
  - egyszerűbb következtetés
    - kvantitatív: mennyi a valószínűsége egy kérdéses megfigyelésnek
    - kvalitatív: mi releváns egy kérdéses vélekedéshez
- A függetlenségek egy eloszlás alapvető kvalitatív jellemzői
  - egy eloszlás összes függetlensége a függetlenségi modellje
  - a függetlenség megfelel az érdektelen/irreleváns hétköznapi fogalmaknak
    - általános szabályszerűségei megfelelnek a "józan észnek": tranzitivitás, monotonitás
    - de ritka kivételek lehetségesek a valószínűségszámítás szerint

Valószínűségi háló (=Bayes-háló)

… egy gráf, ahol

1. Csomópontok: valószínűségi változók egy halmaza
2. Csomópontok között: irányított élek halmaza:
   Az Y csomópontot az X csomóponttal összekötő nyíl $\rightarrow$
   az *Y-nak közvetlen befolyása* van az *X*-re
3. Minden csomópont:
   feltételes valószínűségi tábla $\rightarrow$
   szülők hatása a csomópontra P(X | *Szülők*(X))



4. A gráf nem tartalmaz irányított kört (= **irányított, körmentes gráf = DAG**).

Valószínűségi változó = egy állítás a problémáról

Pl. X = HMG szint magas (két érték: I/N), vagy
X = HMG szint (több érték: Magas / Közepes / Alacsony)

Példa

Otthonunkban egy új **riasztót** szereltünk fel. Ez megbízhatóan észleli a **betöréseket**, de kisebb **földrengés** esetén is jelez.
Két szomszédunk, **János** és **Mária**, megígérték, hogy **felhívnak** a munkahelyünkön, ha meghallják a riasztónkat.
János mindig felhív, ha szól a riasztó, de néha összekeveri a telefoncsörgést a riasztó csengésével és ekkor is telefonál.
Mária viszont, mivel szereti hangosan hallgatni a zenét, néha meg sem hallja a riasztót.
Mi tehát a hívások bekövetkezte vagy hiánya alapján szeretnénk megbecsülni a betörés valószínűségét.

(Bináris) változók (tények):

Betörés (megtörtént): Igen/ Nem

Földrengés (megtörtént): Igen/ Nem

Riasztás (megszólalt): Igen/ Nem

János telefonál(t-e): Igen/ Nem

Mária telefonál(t-e): Igen/ Nem

Háló topológiája = egy absztrakt tudásbázis

általános ok-okozati folyamatokat ír le az adott problémakörben.

A betöréses háló esetén a topológia azt mutatja, hogy a betörés és a földrengés közvetlenül hat a riasztóra, befolyásolja megszólalásának valószínűségét, ellenben János vagy Mária hívásának bekövetkezése csak magán a riasztón múlik – azaz a háló tartalmazza azt a feltevést, hogy ők közvetlenül nem vesznek észre sem a betöréseket, sem a földrengéseket.

Feltételes valószínűségi tábla – FVT

minden egyes csomópontra.

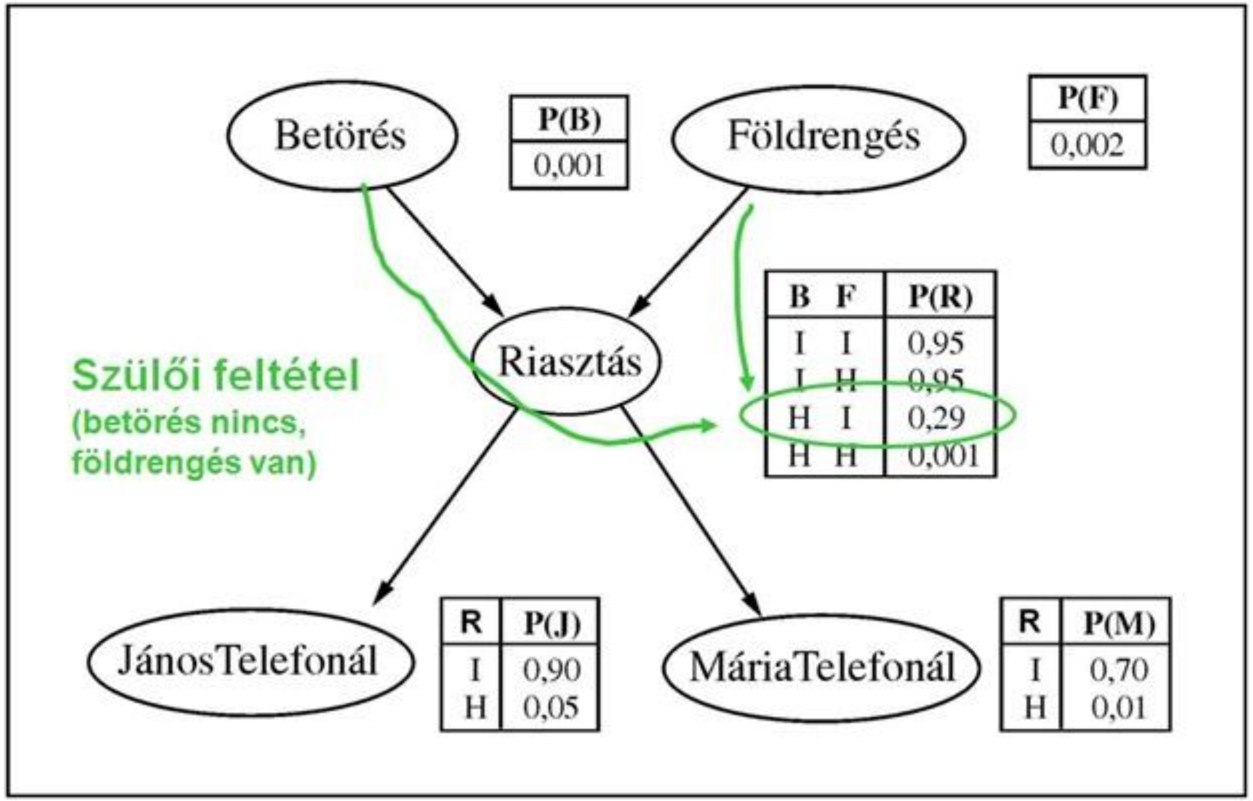
Egy sor a táblázatban az egyes csomóponti értékek feltételes valószínűsége az adott sorhoz tartozó **szülői feltétel** esetén.

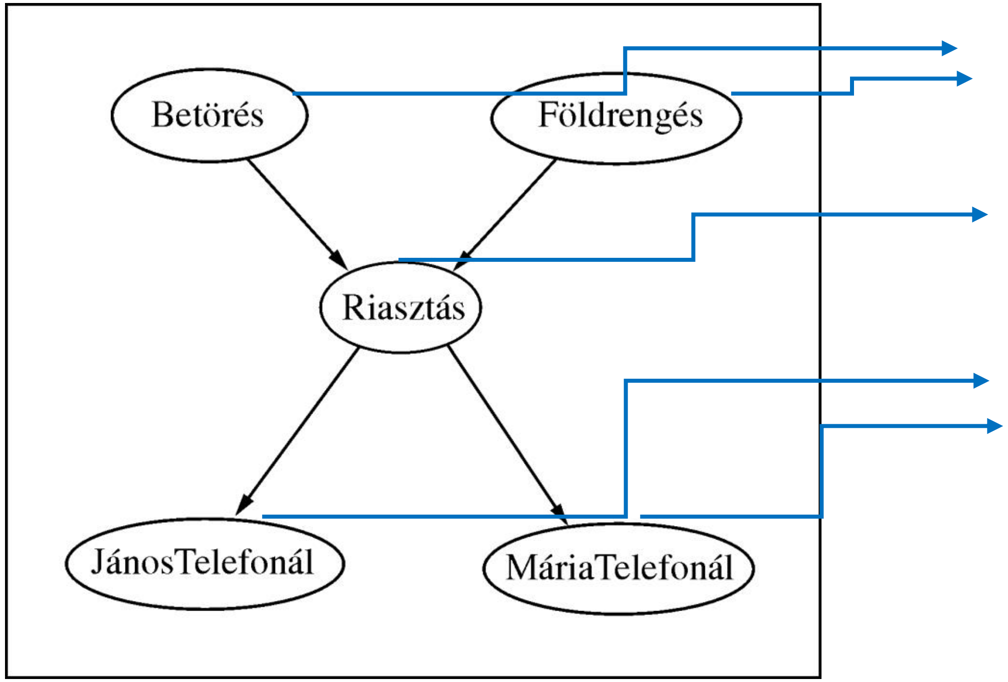
Szülői feltétel: a szülő csomópontok értékeinek egy lehetséges kombinációja (**egyfajta elemi esemény a szülők között**).

Általánosabban, ha a változó **bináris**, és ha *n* **bináris** szülője van:

2^n -1 valószínűség adható meg a feladat alapján.

Szülő nélküli csomópont – a változó egyes értékeinek a priori valószínűségei (ha bináris, akkor csak egy).





Bayes háló:
2 x 1 db a priori
+
1 x 4 db szülői feltétel
+
2 x 2 db szülői feltétel
= 2 + 4 + 4 = **10 db**

Együttes eloszlás:
5 változó
= 2^5-1 = **31 db** valószínűség

Általános eset:

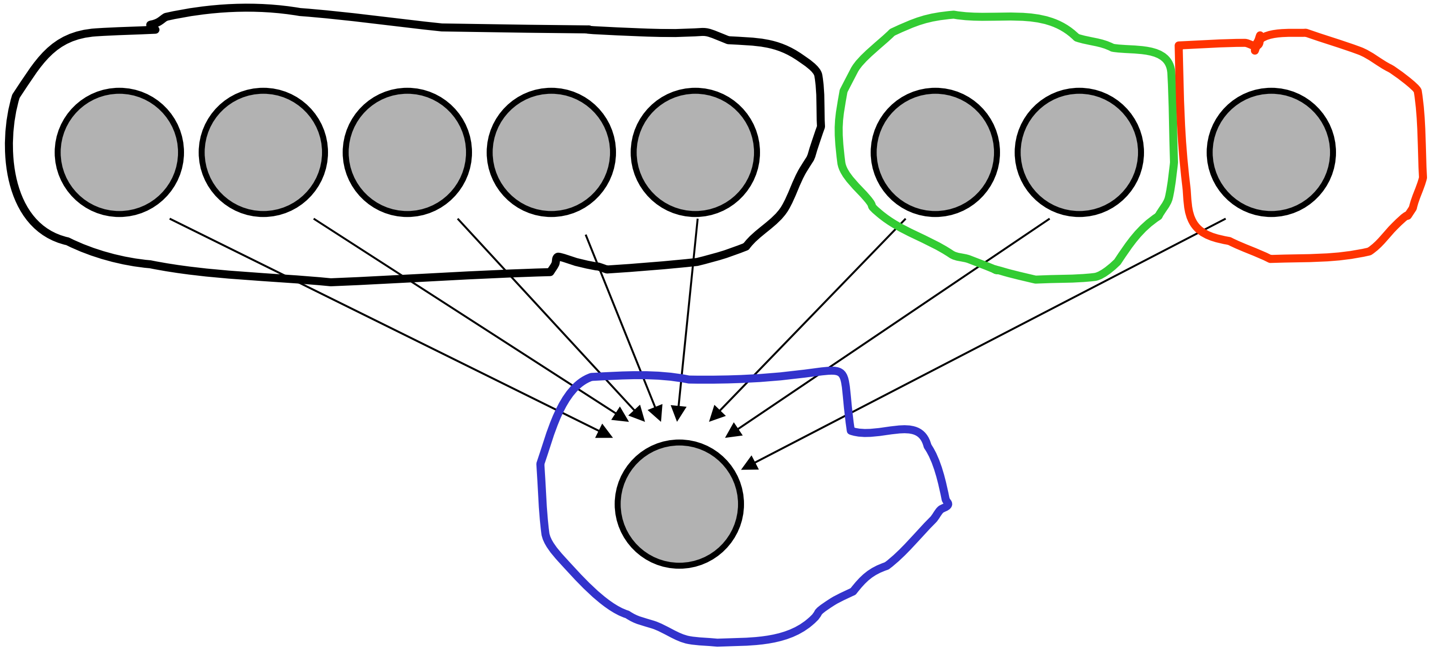
Mi van, pl. ha

5 db szülő csomópont bináris értékű

2 db szülő csomópont 3-as értékű

1 db szülő csomópont 4-es értékű   és

az eredmény csomópont egy 5-ös értékű   változó?



A valószínűségi hálók szemantikája

(1) a háló = **együttes valószínűségi eloszlás** egy leírása,

(2) a háló = **feltételes függetlenségekről szóló állítások** együttese.

A két szemlélet **ekvivalens**
az első abban segít, hogy hogyan **hozzunk létre** egy hálót
a második abban, hogy hogyan **tervezzünk** következtetési eljárásokat.

Feltételes függetlenség

Ha a szülő (szülői feltétel) ismert, nem érdekes az ős:

$\mathbf{P(X_i \mid X_{i-1}, \ldots, \mathit{X_1})} = \boldsymbol{P(X_i \mid \mathit{Szülők}(X_i))}$, i=1…$n$,
ha $\boldsymbol{\mathit{Szülők}(X_i) \subseteq \{ X_{i-1},\ldots,X_1 \}}$.

Ez utóbbi könnyen teljesíthető a csomópontok olyan sorszámozásával, ami konzisztens a gráf implicit részleges rendezésével.

$$\boldsymbol{P(X_i \mid \mathit{Szülők}(X_i), \mathit{Ősök}(X_i)) = P(X_i \mid \mathit{Szülők}(X_i))}$$

Az együttes valószínűségi eloszlásfüggvény leírása

A valószínűségi hálóban található információk alapján:
az együttes valószínűségi eloszlás bármely bejegyzése kiszámítható.

Egy bejegyzés értéke (globális szemantika):

$P(X_1 \ldots X_n) = \prod_{i=1 \ldots n} P(X_i \,|\, Szülők(X_i))$, i=1…$n$

Az együttes valószínűségi eloszlás minden bejegyzése felbontható
a FVT megfelelő elemeinek a szorzatára.
FVT-k: az együttes valószínűségi eloszlás dekomponált leírása.

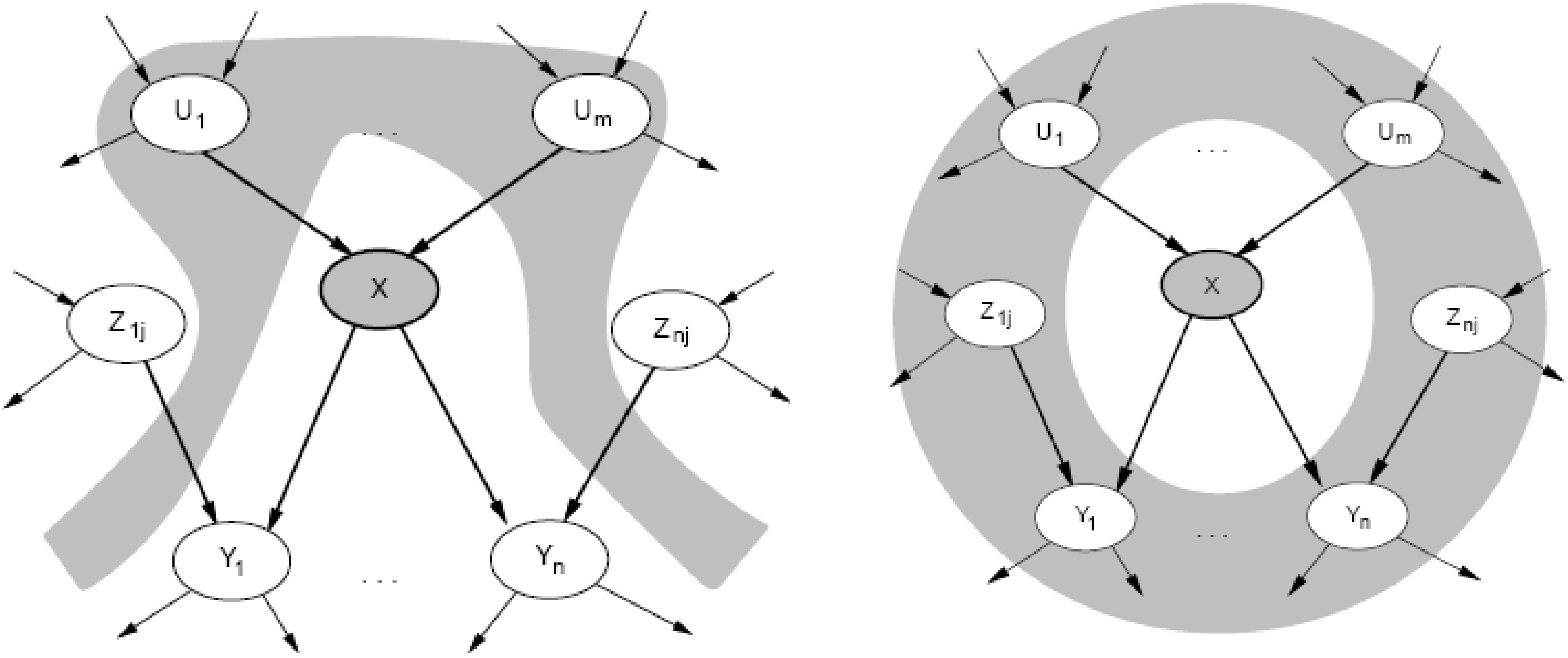
Egy elemi esemény: pl. a riasztó megszólal, de sem betörés,
sem földrengés nem volt, azonban János is és Mária is telefonál:

$$\begin{aligned}
P(J\ M\ R\ \neg B\ \neg F) &= P(J|\ M\ R\ \neg B\ \neg F)\ P(M\ R\ \neg B\ \neg F) \\
&= P(J|R)\ P(M|R\ \neg B\ \neg F)\ P(R\ \neg B\ \neg F) \\
&= P(J|R)\ P(M|R)\ P(R|\neg B\ \neg F)\ P(\neg B)\ P(\neg F) \\
&= .90 \times .70 \times .001 \times .999 \times .998 = .00062
\end{aligned}$$

a. Minden csomópont (X) feltételesen független a nem leszármazottjaitól (Z), ha a **szülői** adottak (U).

b. Minden csomópont (X) feltételesen független minden mástól, ha a **Markov-takarója** adott (szülői+gyerekei+gyerekeinek szülői).

c. **d-elválasztás** …



Az általános eljárás egy háló fokozatos megépítésére

1. Határozzuk meg a problémát leíró változókat (miről érdemes beszélni).

2. Határozzunk meg **egy sorrendet**.

3. Ameddig maradt még érintetlen változó:

   a) Válasszuk a következő X_i változót és adjunk egy csomópontot a hálóhoz.

   b) Legyen a *Szülők*(X_i) a csomópontok azon **minimális halmaza**, amik már szerepelnek a hálóban és a feltételes függetlenség tulajdonságát teljesítik – húzzuk be a kapcsolatokat.

   c) Definiáljuk X_i csomópont feltételes valószínűségi tábláját.

Mindegyik csomópontot csak korábbi csomóponthoz csatlakoztathatunk = a háló körmentes lesz.
A valószínűségi háló nem tartalmaz redundáns valószínűségi értékeket, kivéve soronkénti bejegyzéseket a feltételes valószínűségi táblában.

Tömörség és a csomópontok sorrendje

Egy valószínűségi háló gyakran sokkal **tömörebb**, mint az együttes valószínűségi eloszlásfüggvény.

A valószínűségi háló tömörsége a **lokálisan strukturált** (vagy **ritka**) rendszerek egy példája.

Lokálisan strukturált rendszer: egy komponense csak korlátos számú más komponenssel van kapcsolatban közvetlenül, függetlenül a komponensek teljes számától.

Lokális struktúrák: **inkább a lineáris, mint az exponenciális komplexitás** növekedés.

Lokális struktúrák: **inkább a lineáris, mint az exponenciális komplexitás** növekedés.

Valószínűségi háló n változóból: legtöbb esetben egy változót csak k számú más változó befolyásol, bináris változók = $n \times 2^k$ érték. Együttes valószínűségi eloszlás = $2^n - 1$ érték.

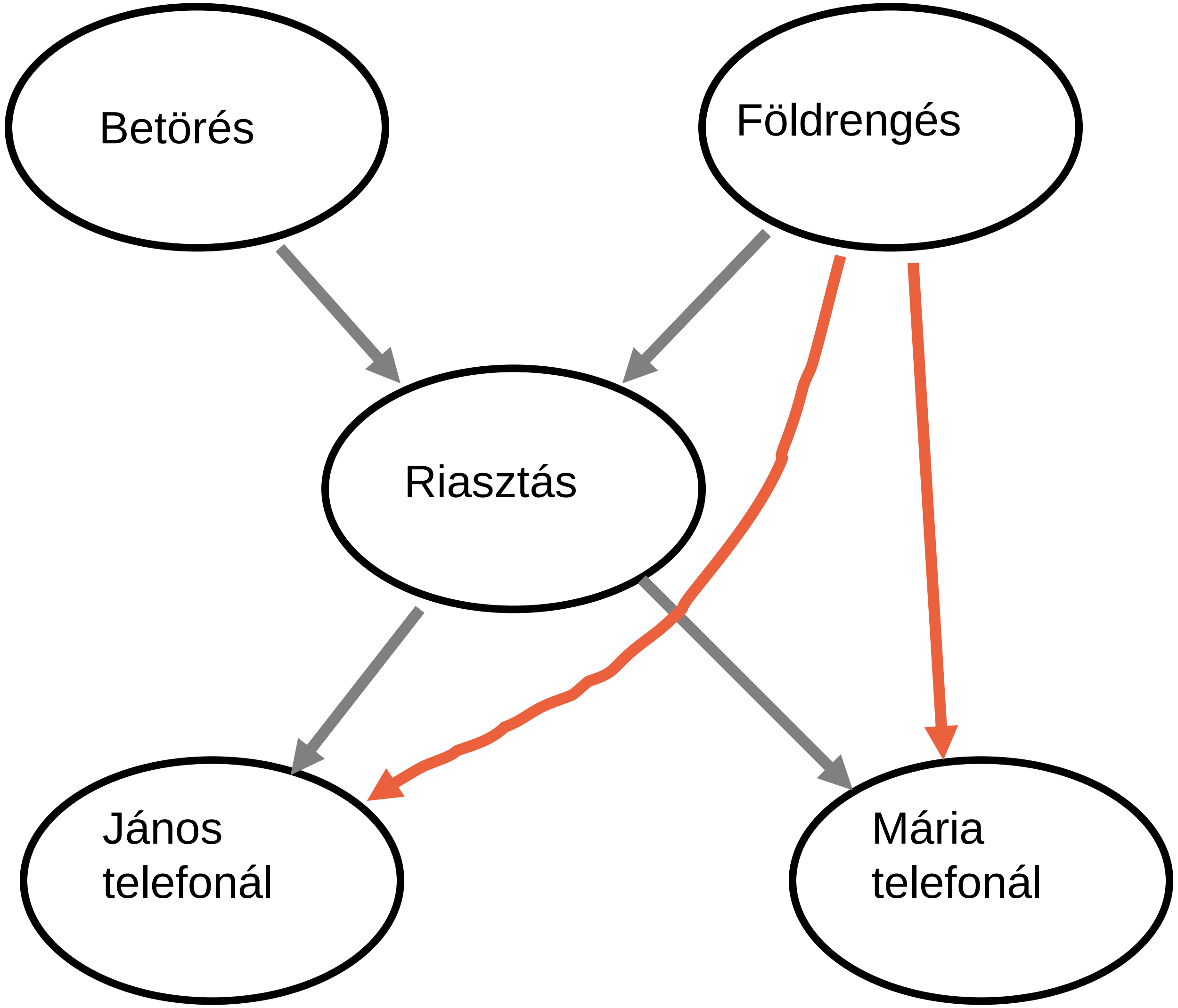
Példa: 20 cs.pont (n = 20), max. 5 szülője legyen (k = 5), valószínűségi háló: 640 érték, együttes valószínűségi eloszlás: > egy millió.

A gyakorlatban előforduló információcsökkenés amiatt lép fel, hogy a valós problémák igen strukturáltak, amit a hálók könnyűszerrel képesek kihasználni.

Tömörség és a csomópontok sorrendje

Bizonyos tárgytartományokban létezhetnek olyan jelentéktelennek tűnő függőségek, amiket feltétlenül modellezni kell egy új kapcsolat felvételével.
De ha ezek a függőségek ténylegesen jelentéktelenek, akkor lehet, hogy nem éri meg a háló komplexitását megnövelni a pontosság kismértékű növelésének érdekében.

Pl.:
hogyha földrengés van, akkor Mária és János akkor sem telefonálna, ha hallanák a riasztót, mivel feltételezik, hogy a földrengés okozta.



2 + 4 + 2 + 2
= 10 (eddig)

2 + 4 + 4 + 4
= 14

Megéri-e?

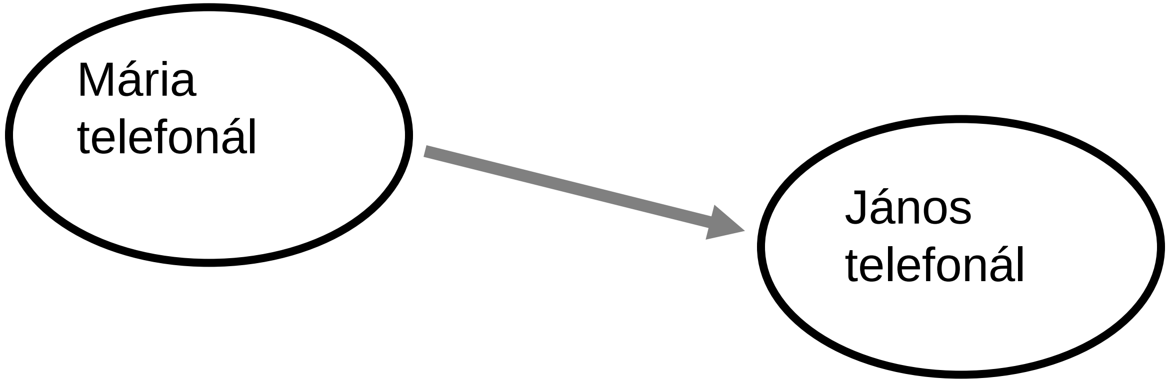
Tömörség és a csomópontok sorrendje

A konstrukciós eljárás működése miatt előbb a „közvetlen befolyásolók”-at kell a hálóhoz adni, ha azt szeretnénk, hogy szülőknek tudjuk őket választani az általuk befolyásolt csomópontnál.

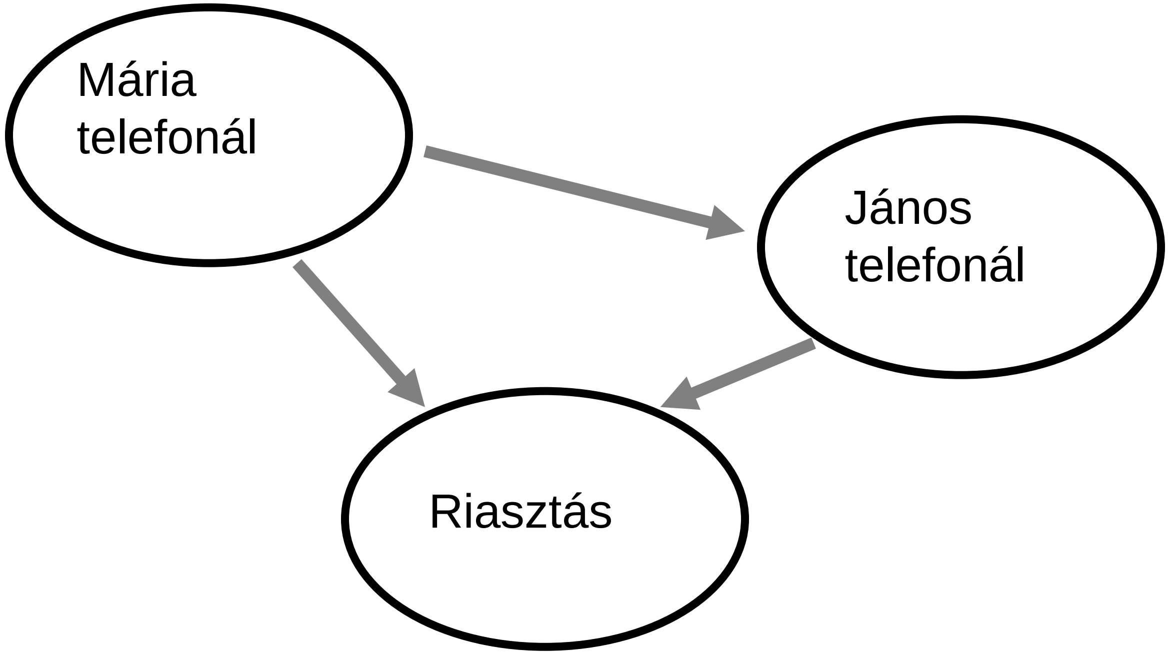
Ezért a helyes sorrend a csomópontok hozzáadásánál:
először az „alapvető okokat” adjuk a hálóhoz, majd a változókat, amiket befolyásolnak, és ezt addig folytatjuk, amíg el nem érjük a „leveleket”, amiknek már nincs közvetlen okozati hatása más változókra (azaz a kauzalitás sorrendjében).

Mi történik, ha történetesen egy rossz sorrendet választunk?
pl.
$MáriaTelefonál \rightarrow JánosTelefonál \rightarrow Riasztás \rightarrow Betörés \rightarrow Földrengés.$

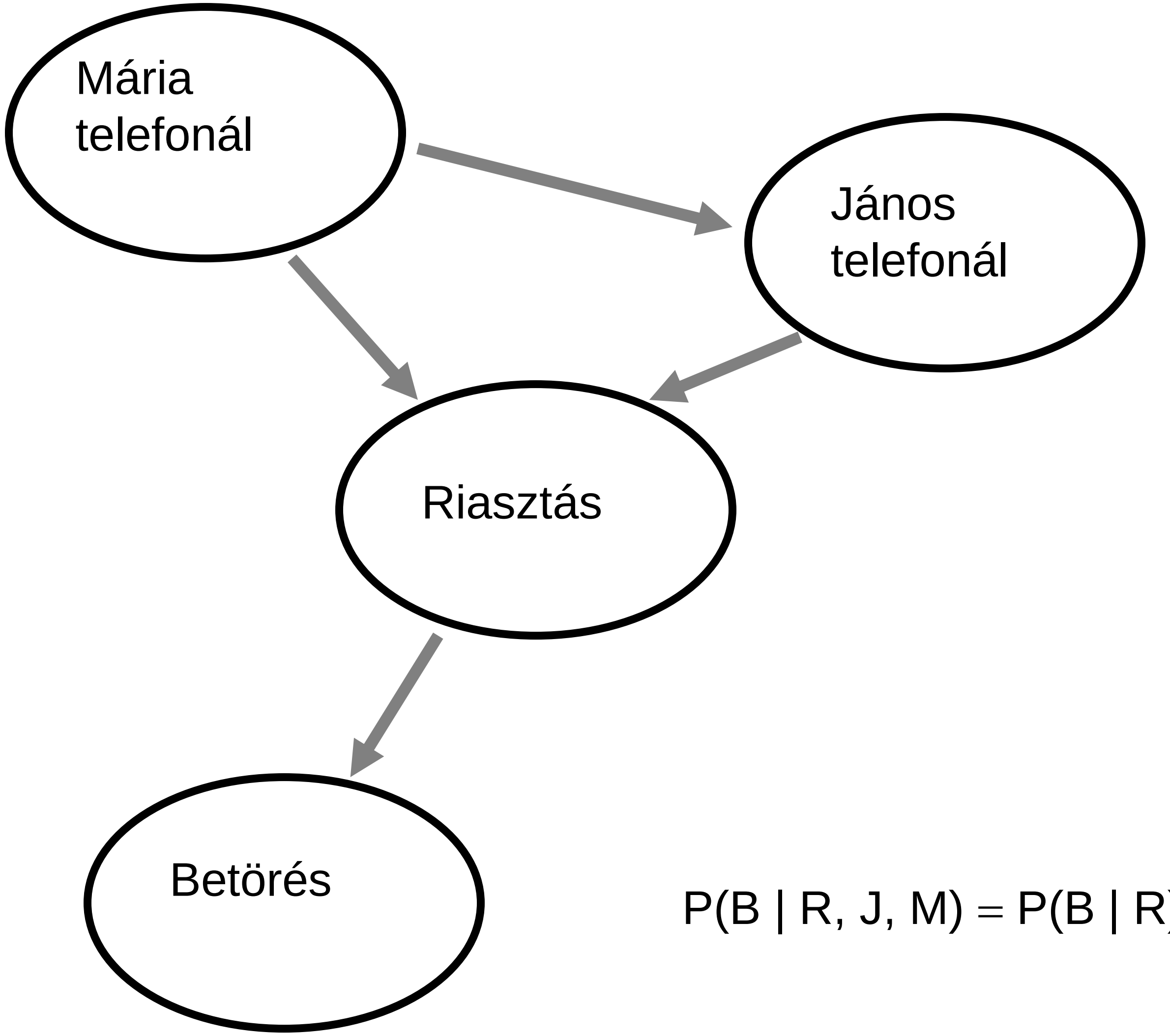
Mária
telefonál



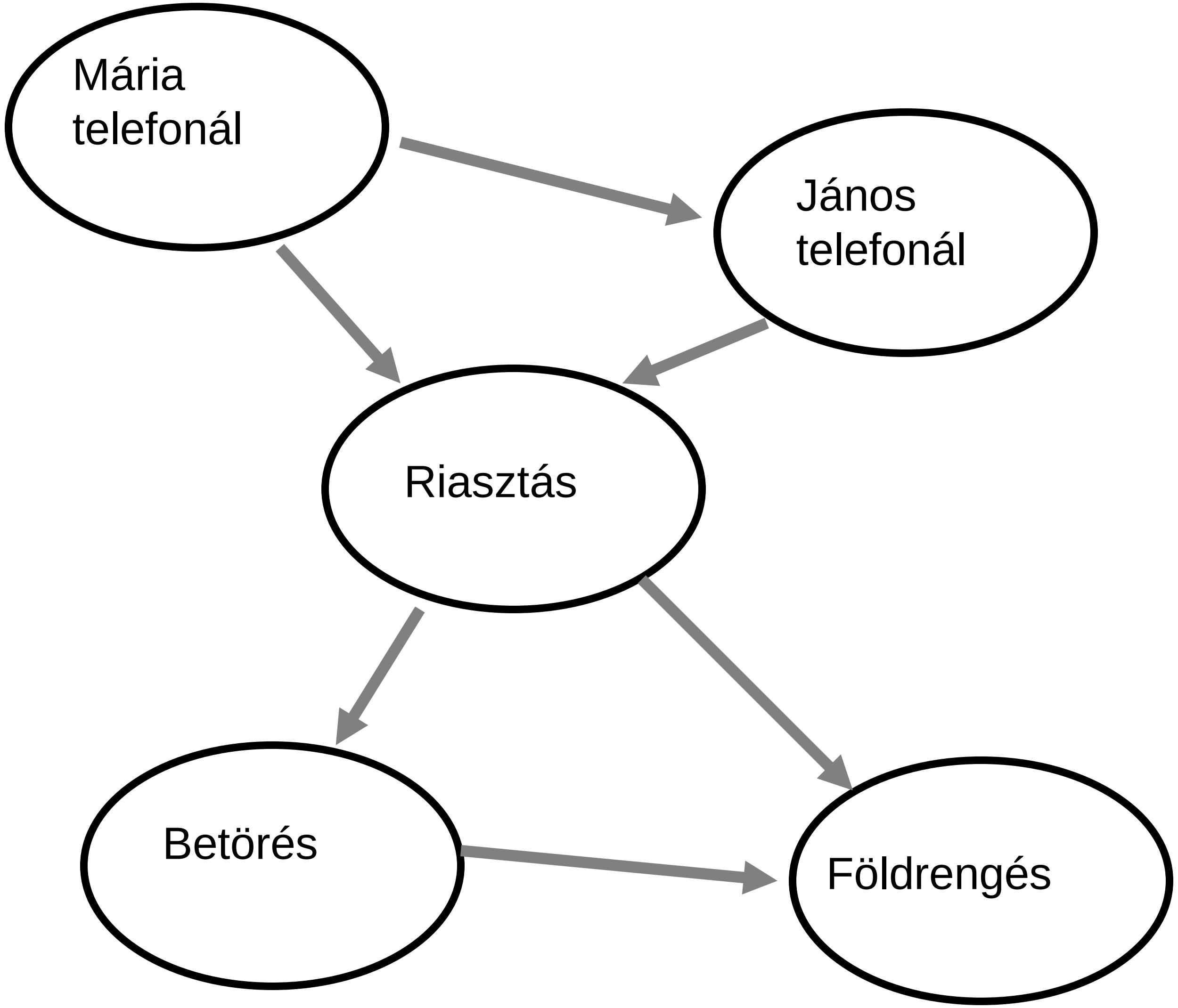
P(J | M) = P(J)?   Nem (azaz van nyíl)



$P(R \mid J, M) = P(R \mid J)$? $P(R \mid J, M) = P(R)$?   Nem



$P(B \mid R, J, M) = P(B \mid R)$?   Igen



1 + 2 + 4
+ 2 + 4
= 13

P(F | B, R, J, M) = P(F | B, R)? = P(F | R)?

Igen Nem

Mi történik, ha történetesen egy nagyon rossz sorrendet választunk?

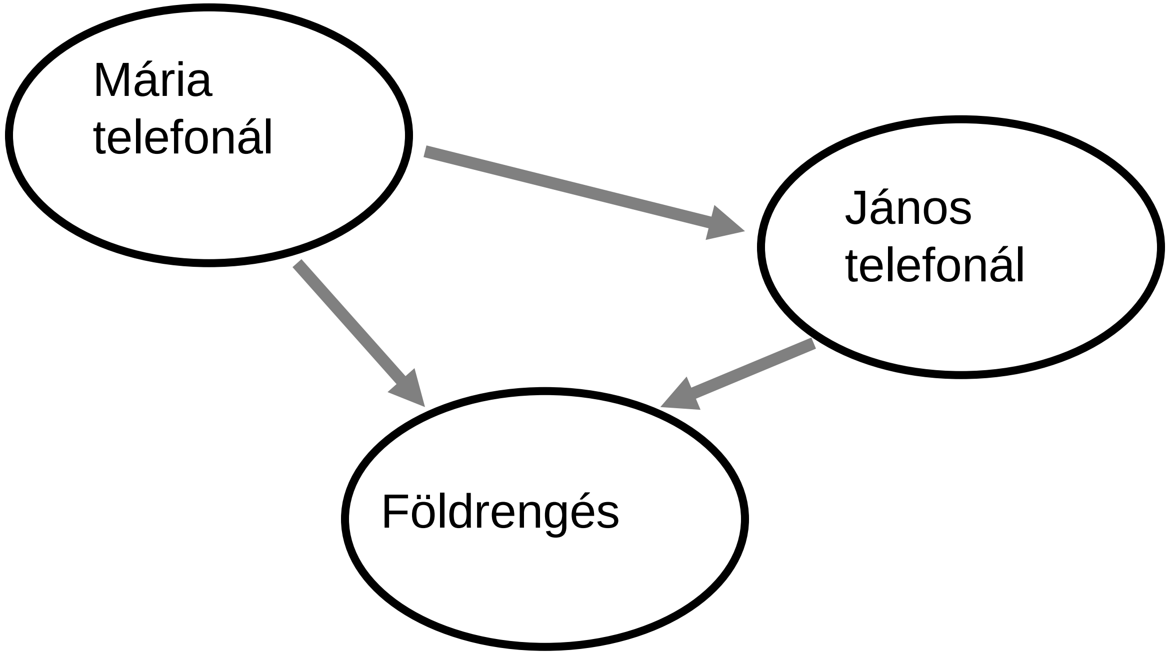
MáriaTelefonál

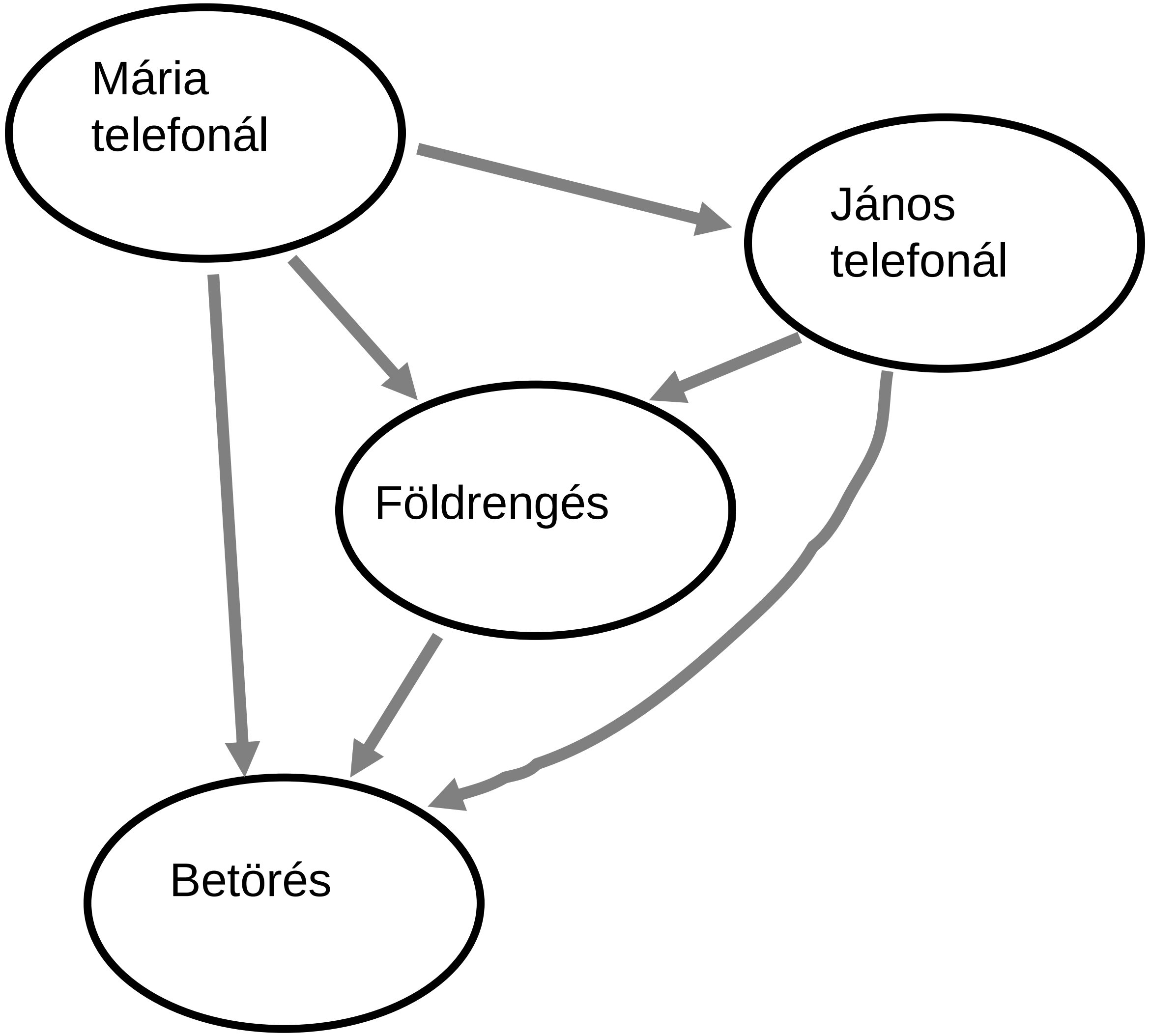
JánosTelefonál

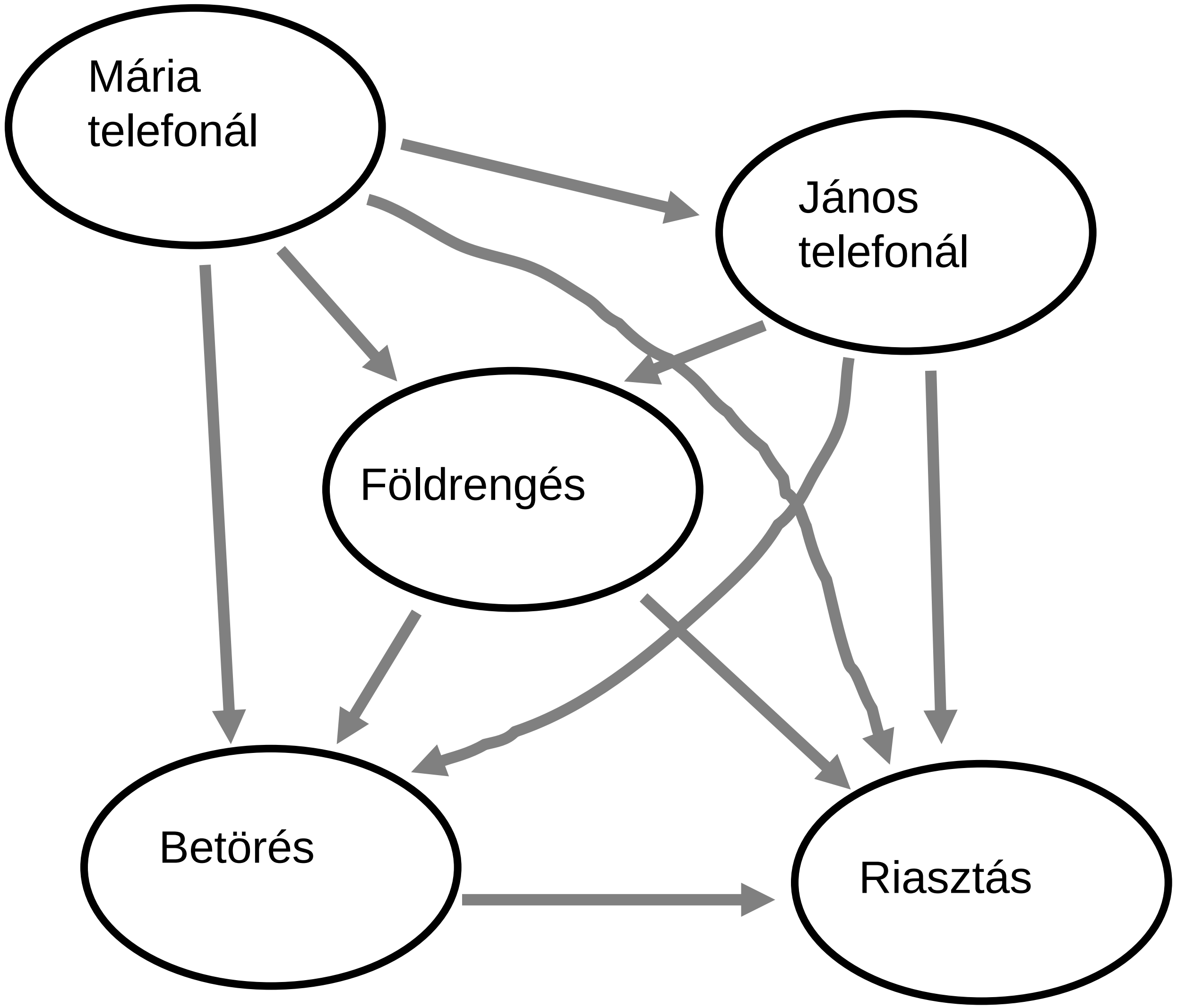
Földrengés.

Riasztás

Betörés


Mária
telefonál
János
telefonál
Földrengés


Mária
telefonál
János
telefonál
Földrengés
Betörés



1 + 2 + 4

+ 8 + 16

= 31

= 2^5 -1

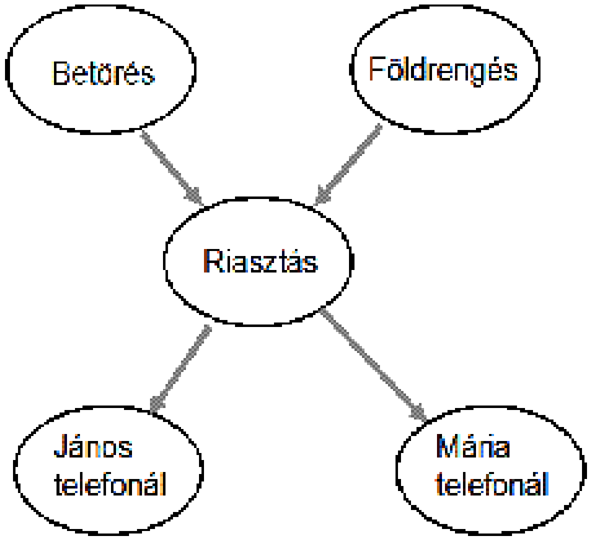
Sok valószínűség (elméletileg exponenciális számú)

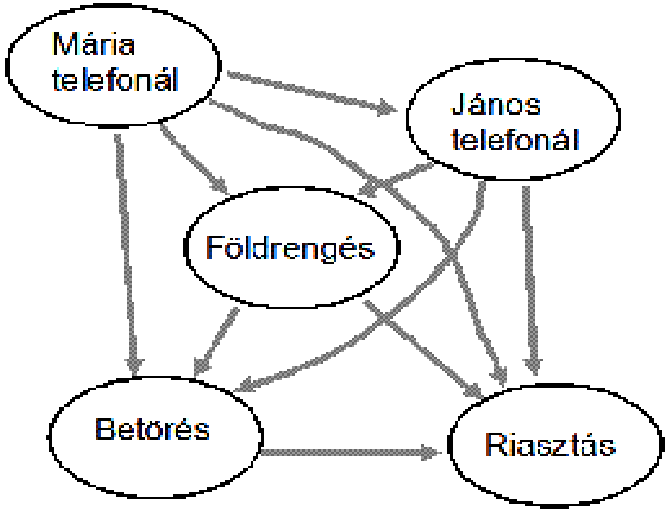
De a legrosszabb következmény:
néhány kapcsolat furcsa viszonyt reprezentál, ami nehéz és **nem természetes valószínűségi ítéleteket igényel**, pl.

P(*Földrengés* | *Mária …, János …*) ?

P(*Betörés* | *Földrengés*) ? ... ...

Okozati (kauzális) modell: kevesebb, megbízhatóbb érték, az értékeket gyakran könnyebb elérni





Következtetés valószínűségi hálókban

Az alapvető feladat:
kiszámítani az a posteriori valószínűséget a **lekérdezéses változók**ra, ha a **tény** ill. **bizonyíték (evidencia) változók**nak az értékei adottak:

P(Lekérdezéses | Bizonyíték)

A riasztós példában, pl.:

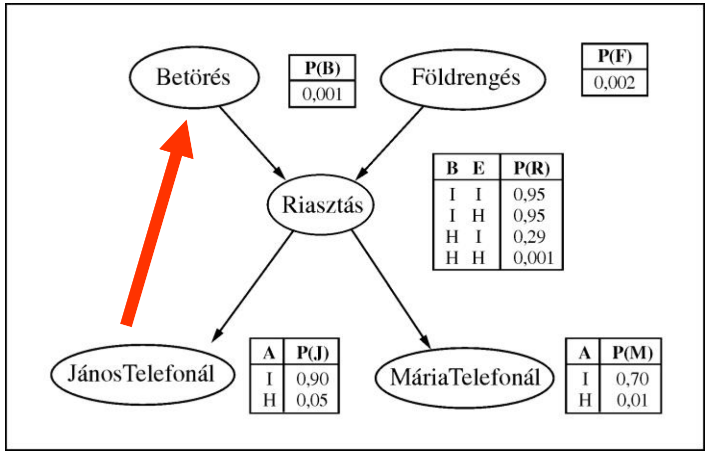
P(*Betörés* | *MáriaTelefonál* és *JánosTelefonál*)

Általában *az ágens az érzékeléséből (vagy egyéb következtetésből) kap értékeket a tény változókhoz, és más változók lehetséges értékeiről kérdez, hogy el tudja dönteni milyen cselekvéseket végezzen*.

Következtetés valószínűségi hálókban

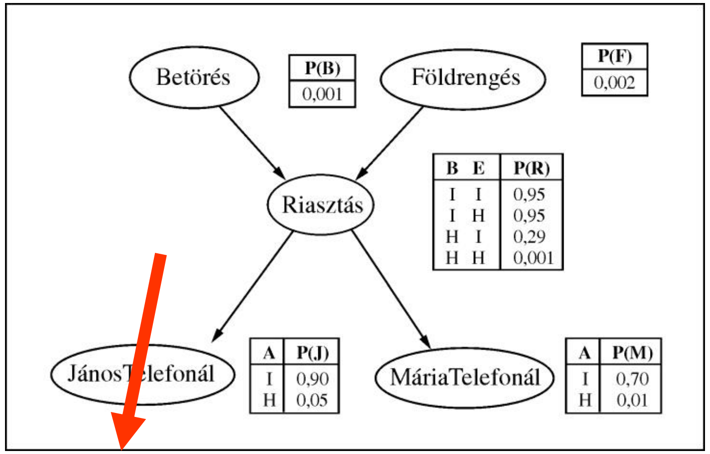
Diagnosztikai következtetés
(hatásról az okra)

Ha adott a *JánosTelefonál,* akkor kiszámíthatjuk pl., hogy
P(*Betörés* | *JánosTelefonál*) = 0,016.



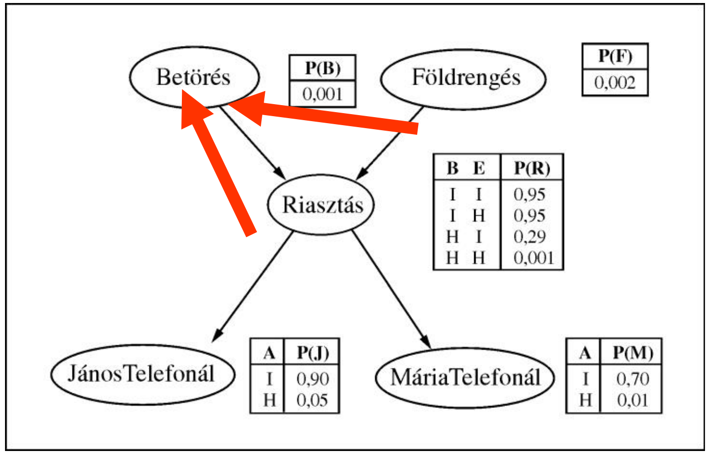
Okozati következtetés
(okról a hatásra)

Ha adott a *Betörés,* akkor kiszámíthatjuk pl., hogy
P(*JánosTelefonál* | *Betörés*) = 0,67.



Okok közötti következtetés

(következtetés egy közös hatás okai között)

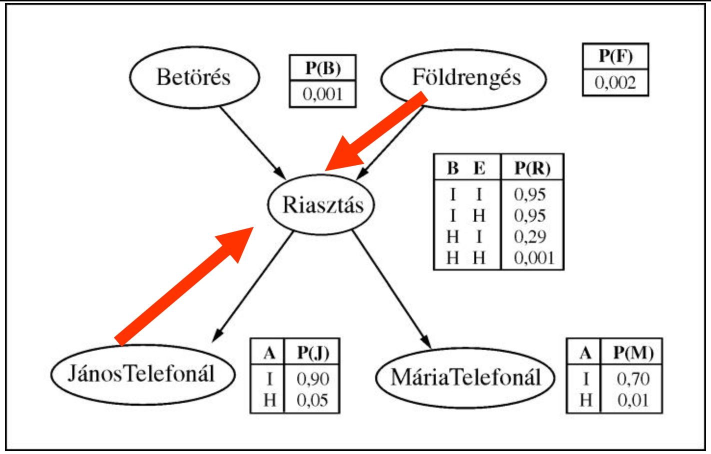


Ha adott a *Riasztás*,
akkor $P(\textit{Betörés} \mid \textit{Riasztás}) = 0.376$.

Ha azonban hozzávesszük azt a tényt is, hogy a *Földrengés* igaz,
akkor $P(\textit{Betörés} \mid \textit{Riasztás}, \textit{Földrengés}) = 0.003$.

Bár a betörések és a földrengések függetlenek, az egyik jelenléte a másik valószínűségét csökkenti.
Ez a fajta következtetési mód a **kimagyarázás.**

Következtetés valószínűségi hálókban



(a fentiek kombinált használata).

Ha a *JánosTelefonál* okozat igaz
és a *Földrengés* ok hamis, akkor

$$P(Riasztás \mid JánosTelefonál, \neg Földrengés) = 0{,}03.$$

Ez a diagnosztikai és okozati következtetés együttes felhasználása.

Hasonlóan,

$$P(Betörés \mid JánosTelefonál, \neg Földrengés) = 0{,}017.$$

Ez a diagnosztikai és az okok közötti következtetés kombinálása.

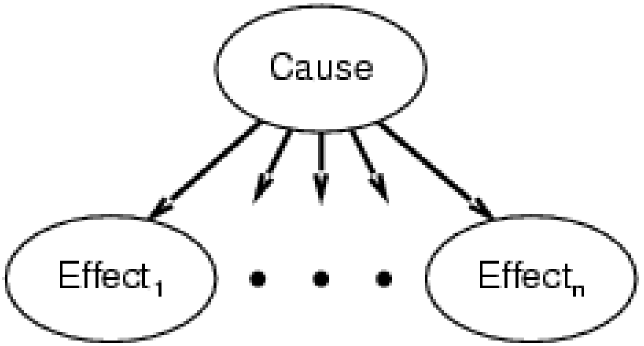
Érzékenységi vizsgálat elvégzése, annak érdekében,
hogy megértsük, hogy a modell mely vonatkozásának van
a legnagyobb hatása a lekérdezéses változókra nézve.

Naiv Bayes-hálók

Feltevések:

1.) Kétféle csomópont lehetséges: a „ok” és „következmény”.

2.) A következmények egymástól feltételesen függetlenek egymástól feltéve az okot.



Naiv Bayes-hálók

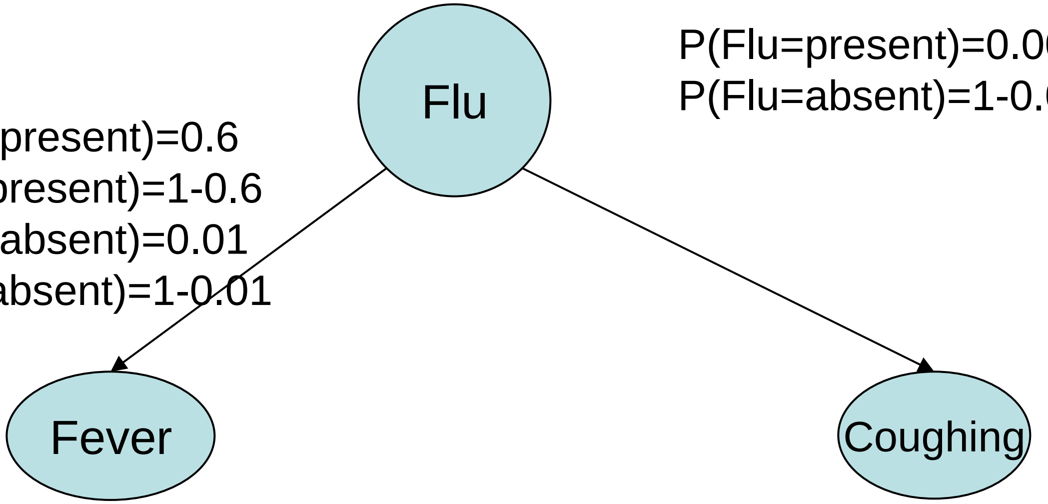
Változók (csomópontok)

Flu (Influenza): {jelen, nincs jelen}
Fever (Láz): {jelen, nincs jelen}
Coughing (Köhögés): {jelen, nincs jelen}

Modell

P(Flu=present)=0.001
P(Flu=absent)=1-0.001

P(Fever=present|Flu=present)=0.6
P(Fever=absent|Flu=present)=1-0.6
P(Fever=present|Flu=absent)=0.01
P(Fever=absent|Flu=absent)=1-0.01



P(Coughing=present|Flu=present)=0.3
P(Coughing=absent|Flu=present)=1-0.3
P(Coughing=present|Flu=absent)=0.02
P(Coughing=absent|Flu=absent)=1-0.02

Naiv Bayes-hálók

Együttes valószínűség-eloszlás dekompozíciója:

$P(Y,X_1,..,X_n) = P(Y)\prod_i P(X_i,|Y, X_1,..,X_{i-1})$ //lánc szabály miatt

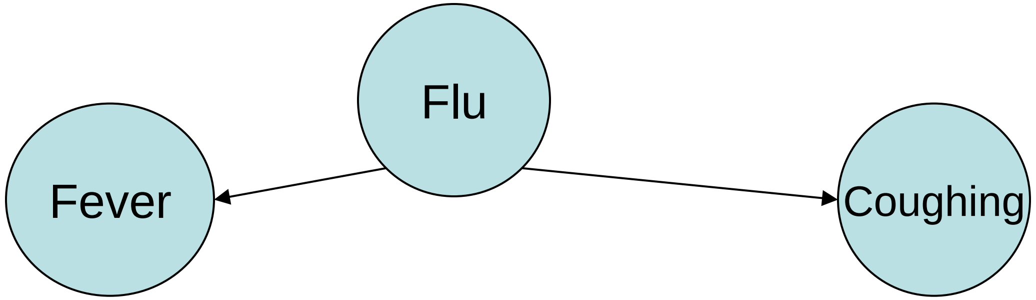
$= P(Y)\prod_i P(X_i,|Y)$ // naiv BN feltevés

2n+1 paraméter

Diagnosztikus következtetés:

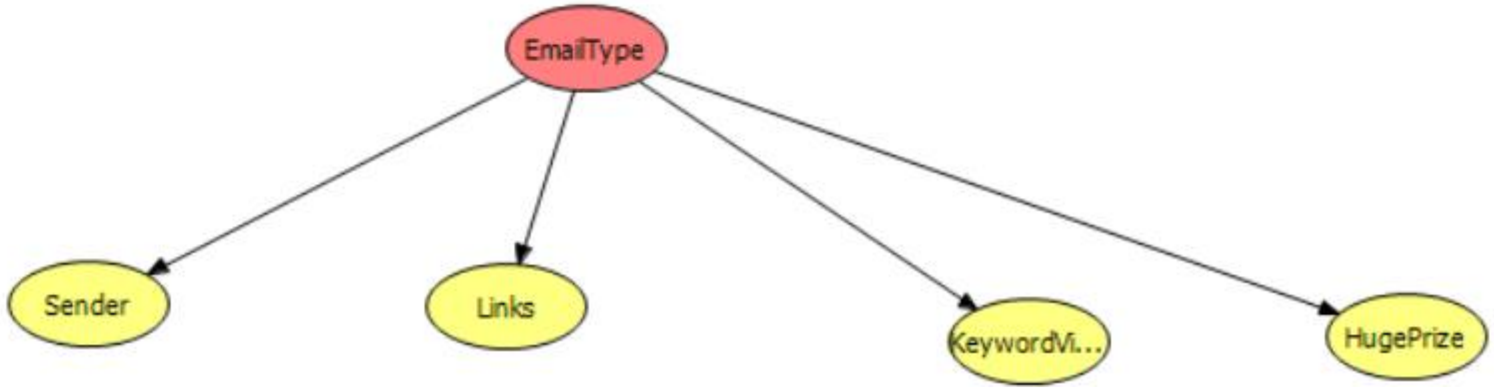
$P(Y|x_{i1},..,x_{ik}) = P(Y)\prod_j P(x_{ij},|Y) / P(x_{i1},..,x_{ik})$

$$p(Flu = present \mid Fever = absent, Coughing = present)$$
$$\propto p(Flu = present)\, p(Fever = absent \mid Flu = present)\, p(Coughing = present \mid Flu = present)$$



Gyakorlati példa: SPAM filter

- SPAM filter
  - SPAM: yes/no [suspicious..]
  - Attributes
    - Sender, subject, link, attachment,..



A valószínűségi hálók szemantikája

(1) a háló = **együttes valószínűségi eloszlás** egy leírása,

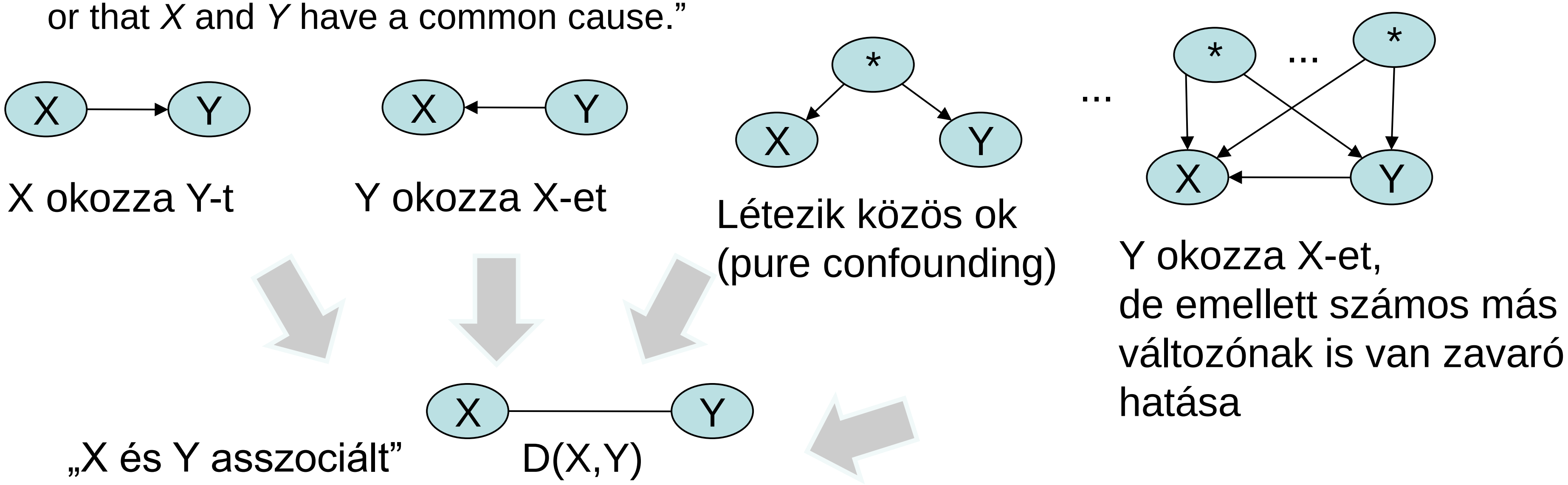
(2) a háló = **feltételes függetlenségekről szóló állítások** együttese.

(3) a háló = **oksági kapcsolatok** együttese.

FONTOS: oksági kapcsolat ≠ asszociációs kapcsolat

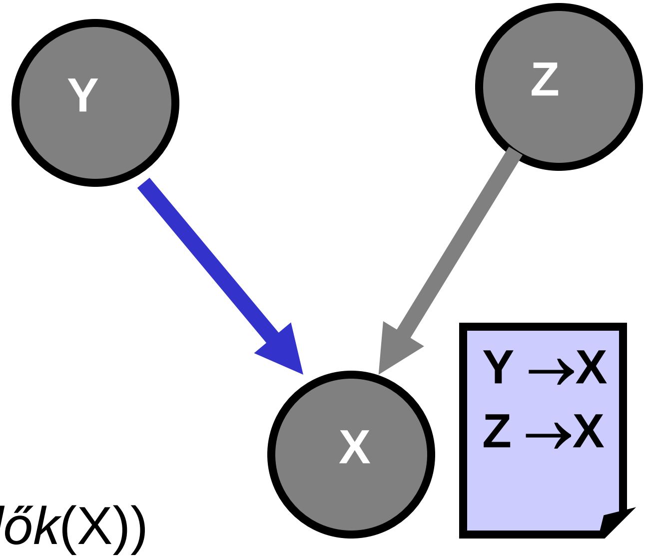
Reichenbach's Common Cause Principle:

„a correlation between events *X* and *Y* indicates either that *X* causes *Y*, or that *Y* causes *X*, or that *X* and *Y* have a common cause."



Valószínűségi háló (Bayes-háló) → egy gráf

1. Csomópontok: valószínűségi változók egy halmaza
2. Csomópontok között: irányított élek halmaza:
   Az Y csomópontot az X csomópponttal összekötő nyíl →
   az *Y-nak közvetlen befolyása* van az *X*-re
3. Minden csomópont:
   feltételes valószínűségi tábla →
   szülők hatása a csomópontra P(X | *Szülők*(X))



4. A gráf nem tartalmaz irányított kört (= **irányított, körmentes gráf = DAG**).

Valószínűségi változó = egy állítás a problémáról

Pl. X = HMG szint magas (két érték: I/N), vagy
X = HMG szint (több érték: Magas / Közepes / Alacsony)