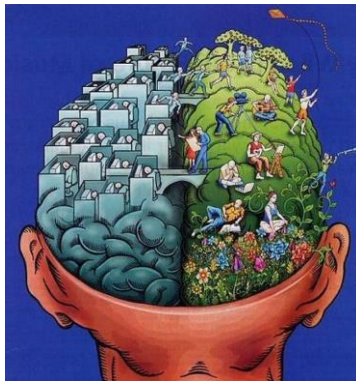




Budapesti Műszaki és Gazdaságtudományi Egyetem Méréstechnika és Információs rendszerek Tanszék



Mesterséges Intelligencia - MI Valószínűségi hálók

Előadó:

Hullám Gábor
Pataki Béla

Előadás anyaga:

Dobrowiecki Tadeusz



BME I.E. 414, 463-26-79

pataki@mit.bme.hu

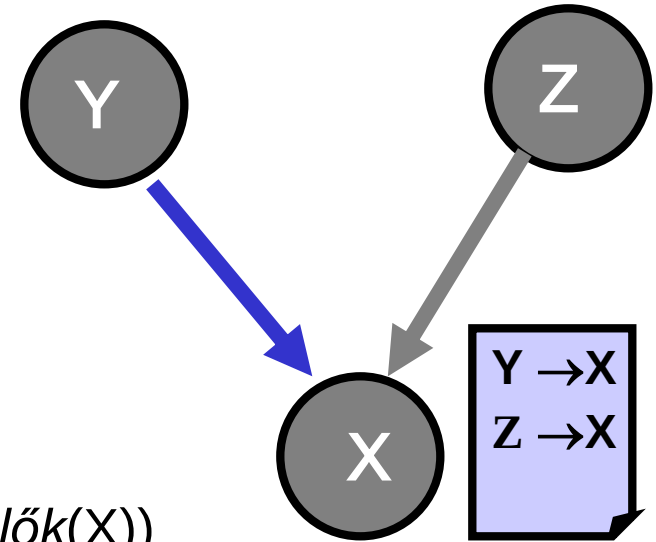
<http://www.mit.bme.hu/general/staff/pataki>



Valószínűségi háló (Bayes-háló) → egy gráf

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

$$P(X \mid \text{Szülők}(X))$$



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

sok esetben érvényes marad, mivel **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.

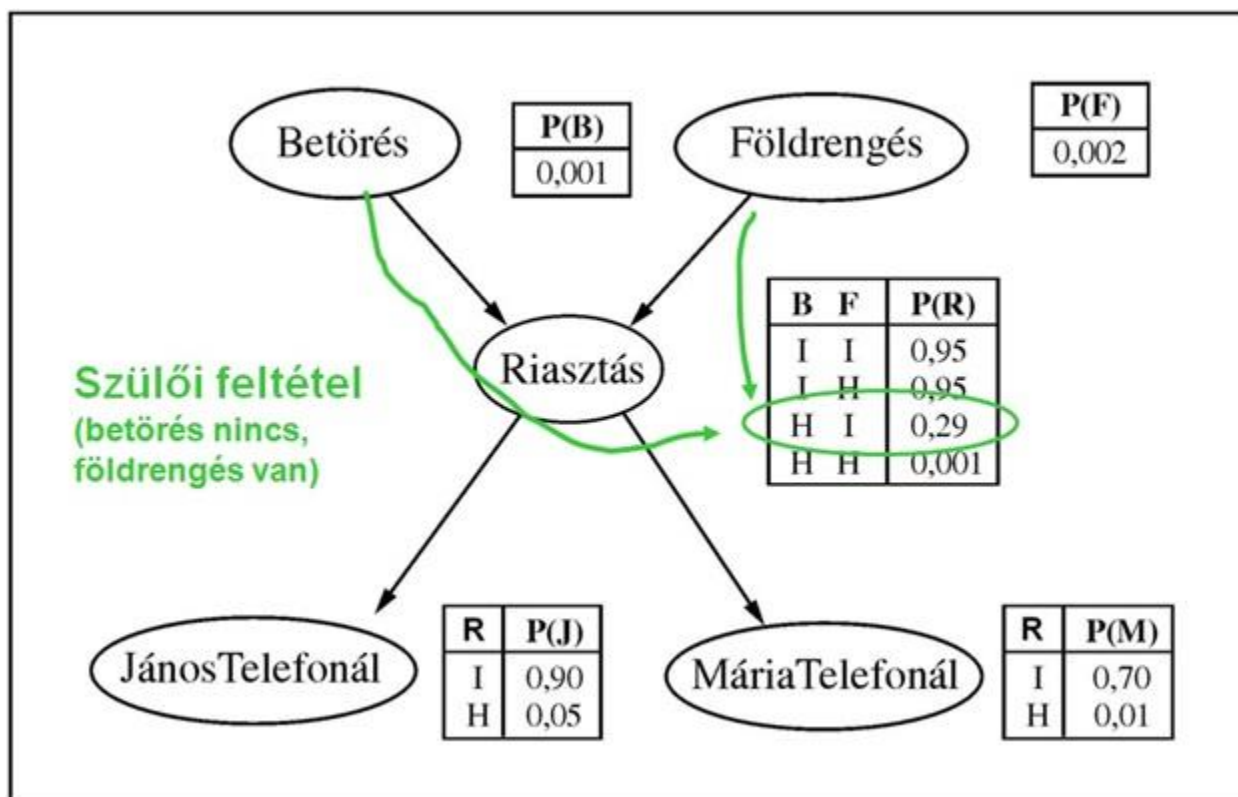
Egy sor a táblázatban az egyes csomóponti értékek feltételes valószínűsége az adott sorhoz tartozó **szülői feltétel** esetén.

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.

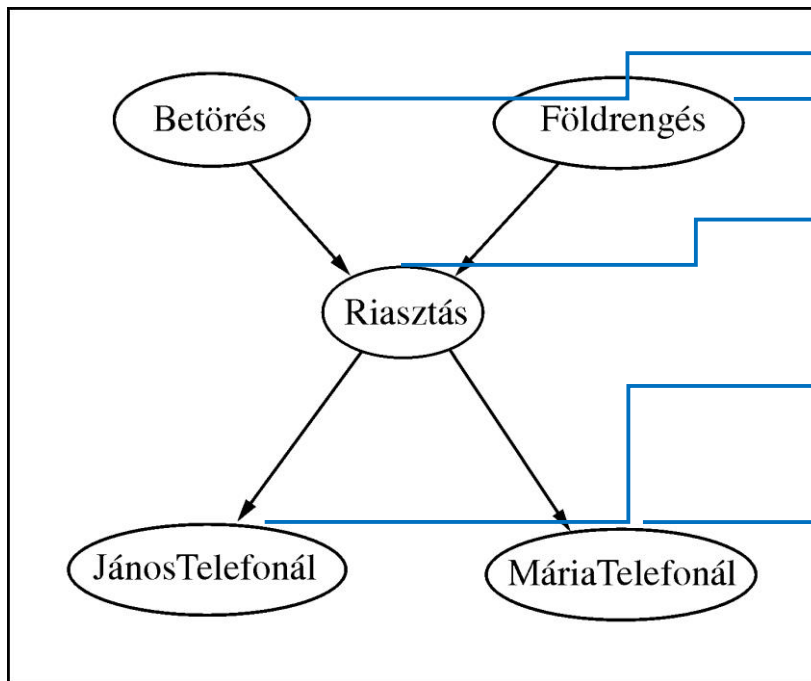
Szülő nélküli csomópont – a változó egyes értékeinek a priori valószínűségei (ha bináris, akkor csak egy).



Együttes eloszlás:

5 változó

= $2^5 - 1 = 31$ db valószínűség



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

Elég-e?

Általános eset:

Mi van, pl. ha

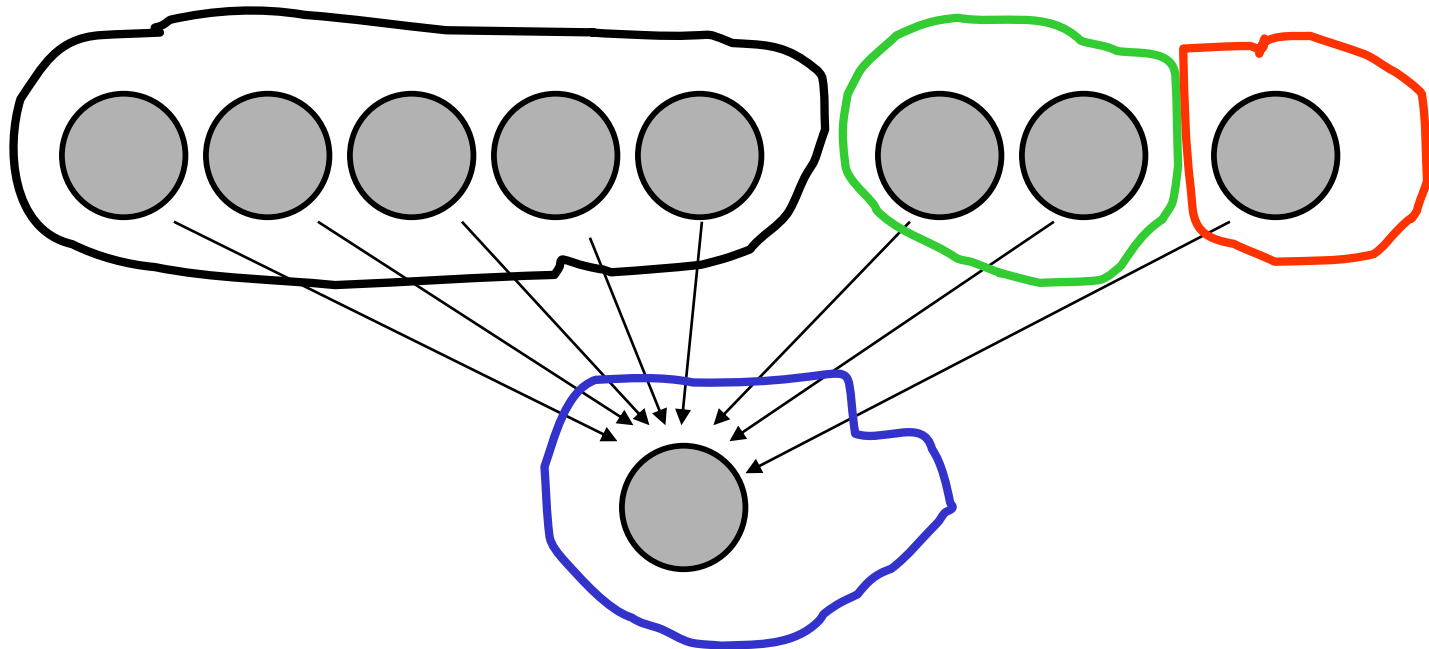
5 db szülő csomópont bináris értékű

2 db szülő csomópont 3-as értékű

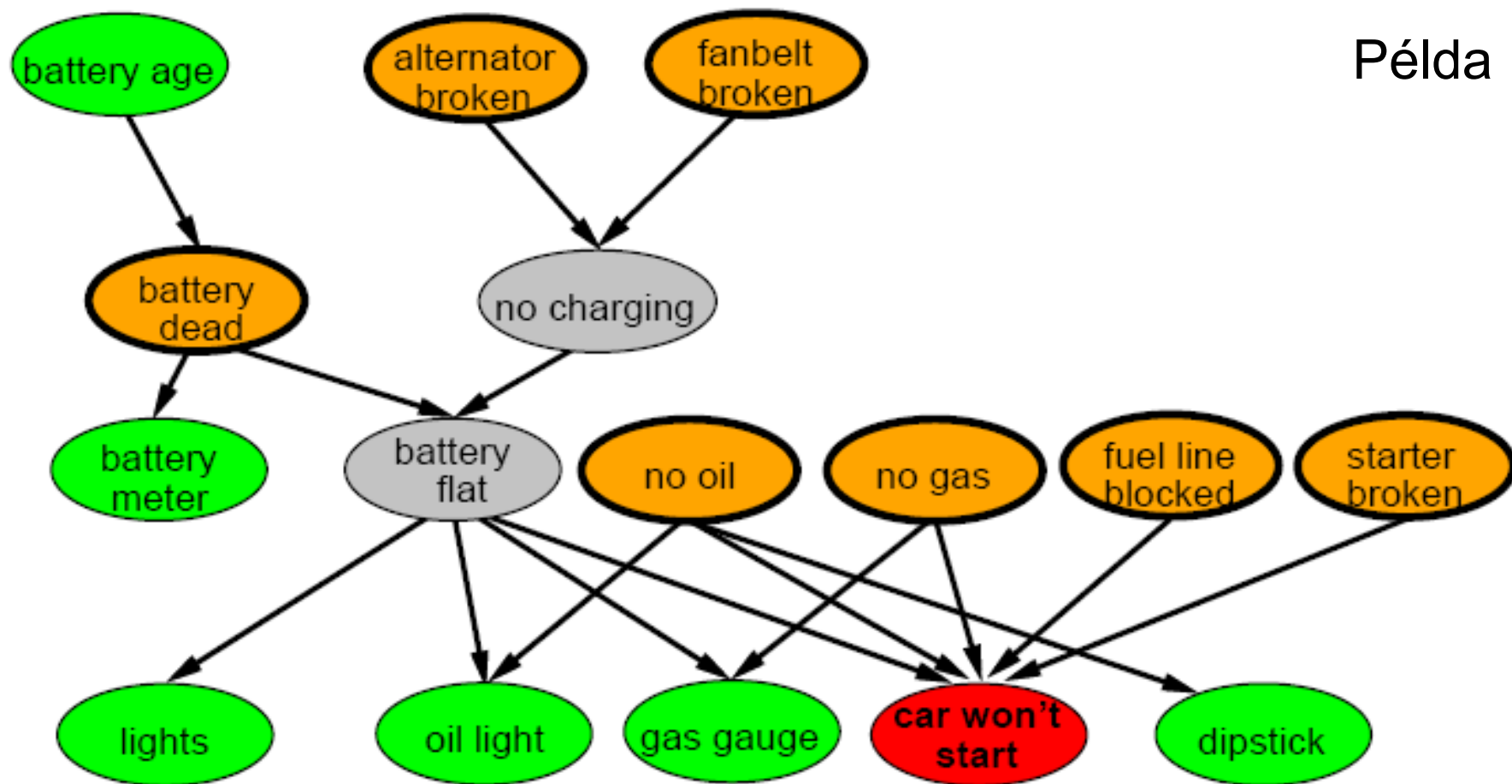
1 db szülő csomópont 4-es értékű és

az eredmény csomópont egy 5-ös értékű

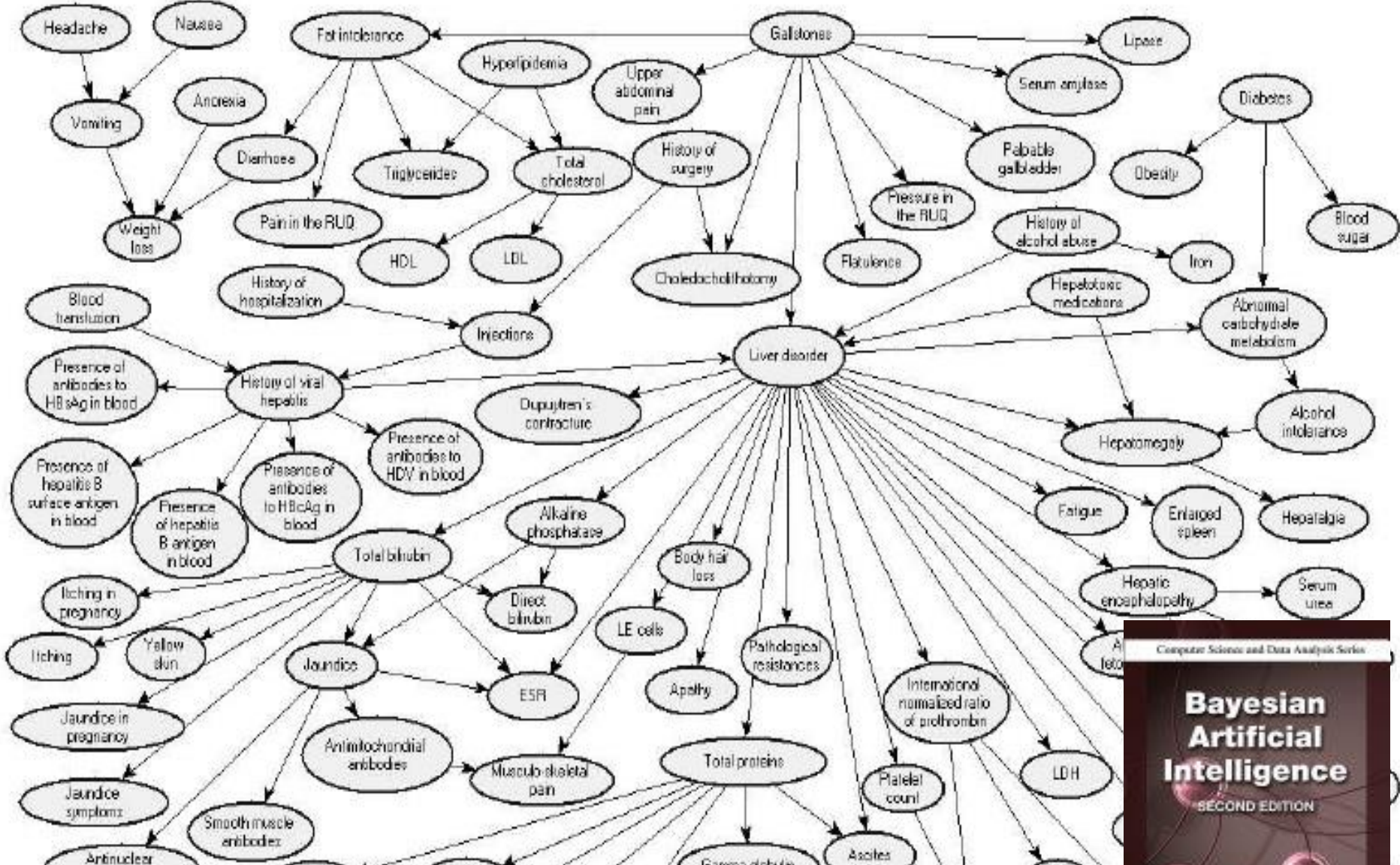
változó?



Példa

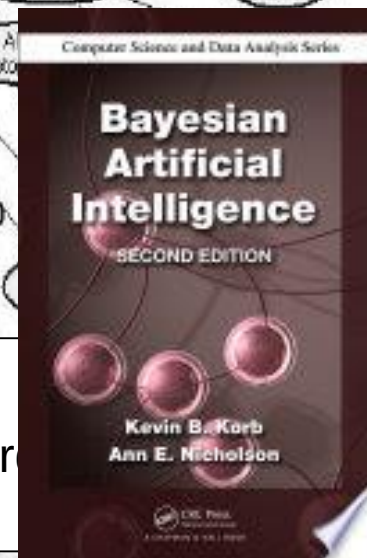


16 (bináris) változó = $2^{16}-1 = 16383$ valószínűség
háló = $7 + 4 \times 2 + 4 \times 4 + 1 \times 32 = 63$ valószínűség (x 260)



Májdiagnózis

A. Onisko et al., Extension of the HEPAR II Model to Multiple-Disorder Diagnosis (2000)



Biológiai és környezeti folyamatok modellezése

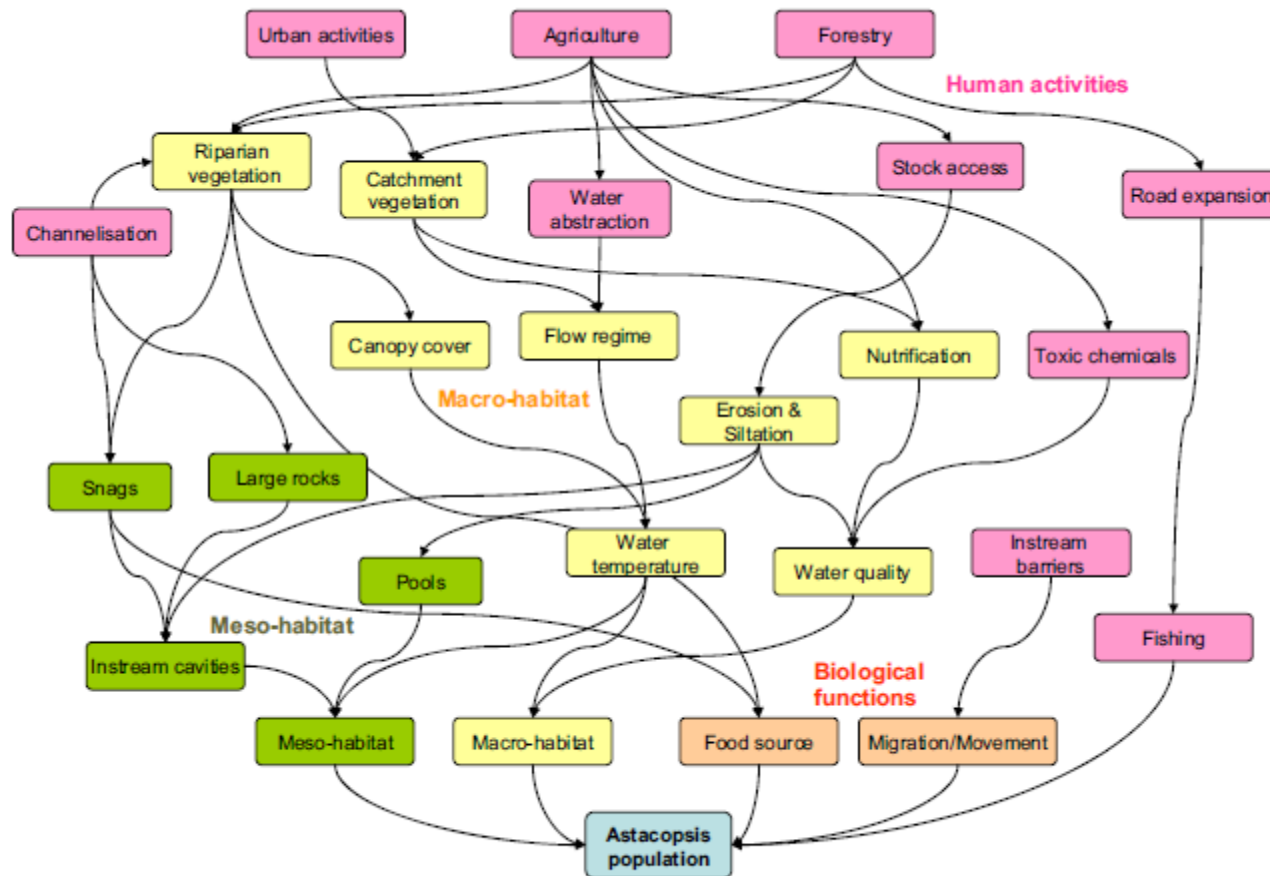
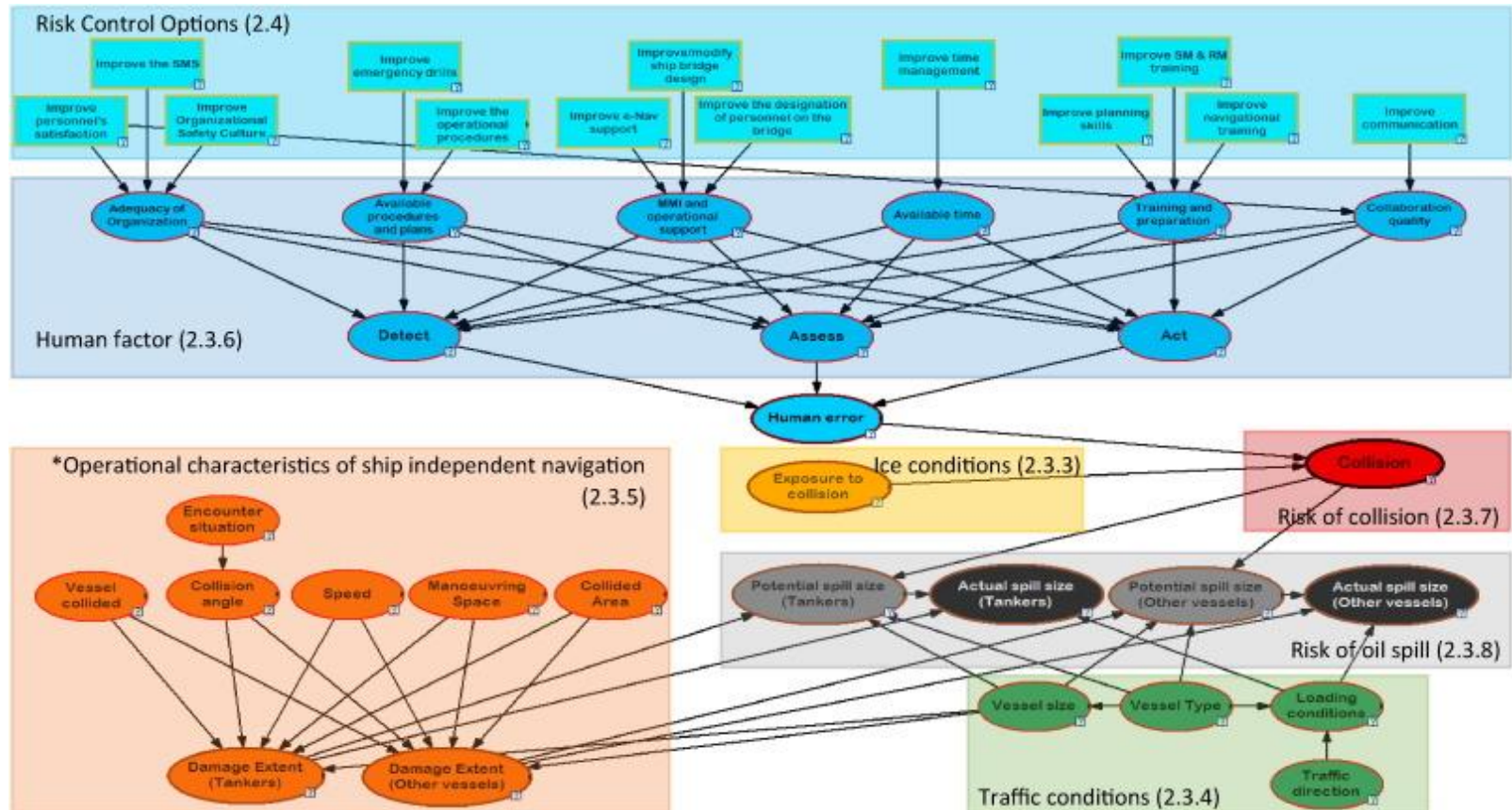


Fig. 2. Conceptual model of the key factors affecting *Astacopsis gouldi* populations.

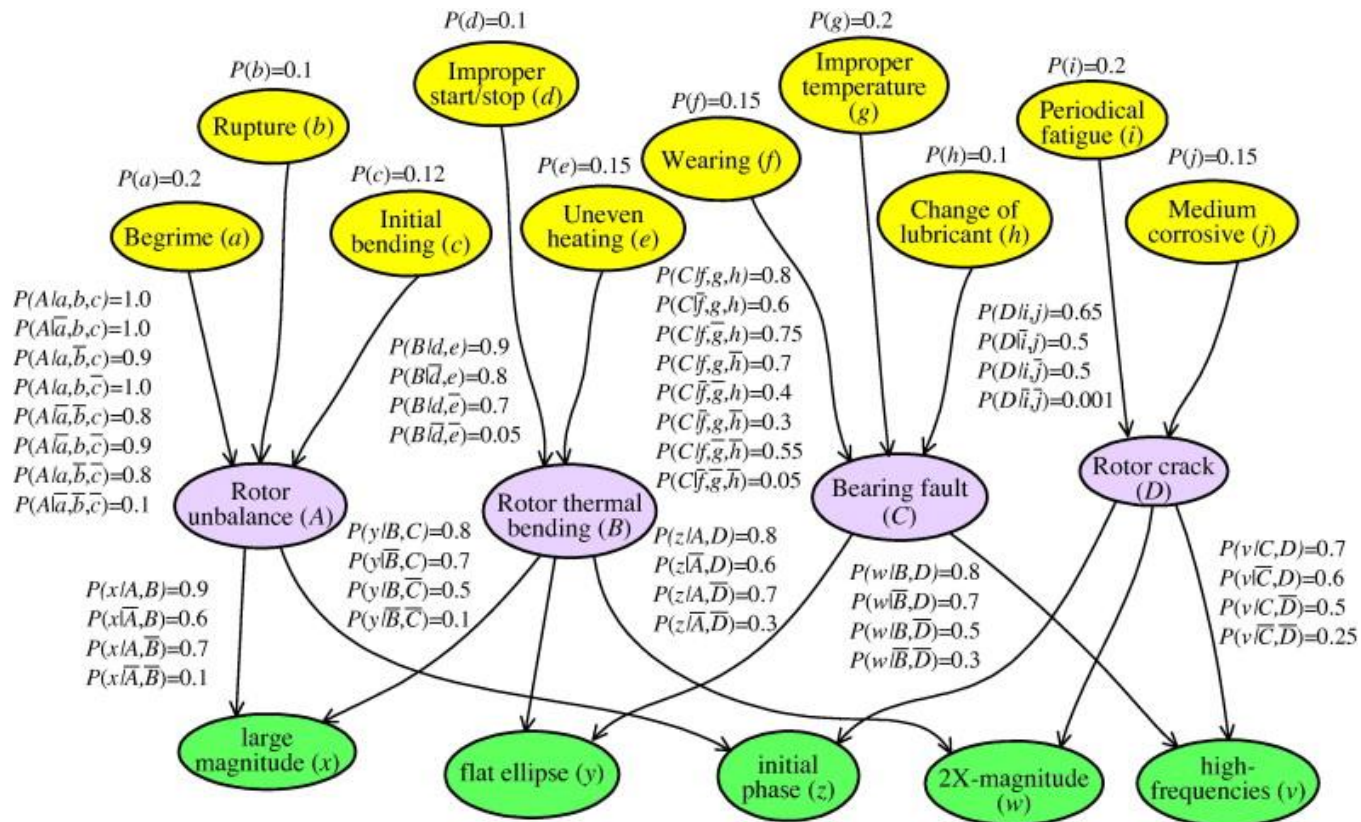
Serena H. Chen & Carmel A. Pollino: Good practice in Bayesian network modelling

Kockázatmodellezés



Osiris A.Valdez et al. : Risk management model of winter navigation operations, Marine Pollution Bulletin, Volume 108, Issues 1–2, 15 July 2016, Pages 242-262

Hibamodellelés és következtetés

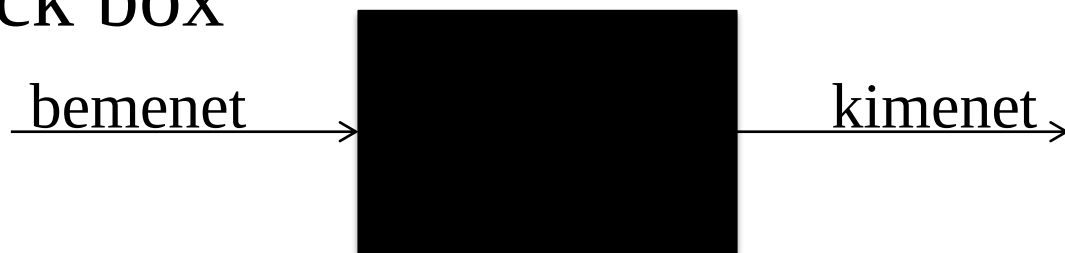


Bin GangXu: Intelligent fault inference for rotating flexible rotors using Bayesian belief network, Expert Systems with Applications, Volume 39, Issue 1, January 2012, Pages 816-822

Modelltípusok

egy rendszer belső működésének modellezése szerint

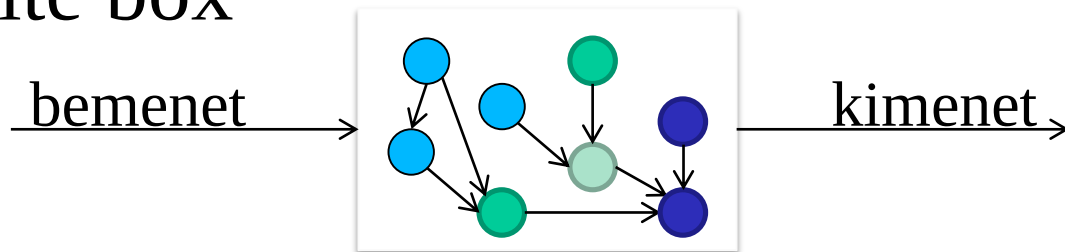
- Black box



Pl.: Neurális hálók

A rendszer belső működését nem modellezi

- White box



Pl.: Bayes- hálók

A rendszer belső működését modellezi

(Grey box)

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 ősz:

$$P(X_i \mid X_{i-1}, \dots, X_1) = P(X_i \mid \mathbf{Szülők}(X_i)), \quad i=1 \dots n,$$

ha $\mathbf{Szülők}(X_i) \subseteq \{X_{i-1}, \dots, X_1\}$.

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

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

Az együttes valószínűségi eloszlásfüggvény leírása

A valószínűségi hálóban található információk alapján:

az együttes valószínűségi eloszlás bármely bejegyzése kiszámítható.

Egy bejegyzés értéke (globális szemantika):

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

Az együttes valószínűségi eloszlás minden bejegyzése felbontható a FVT megfelelő elemeinek a szorzatára.

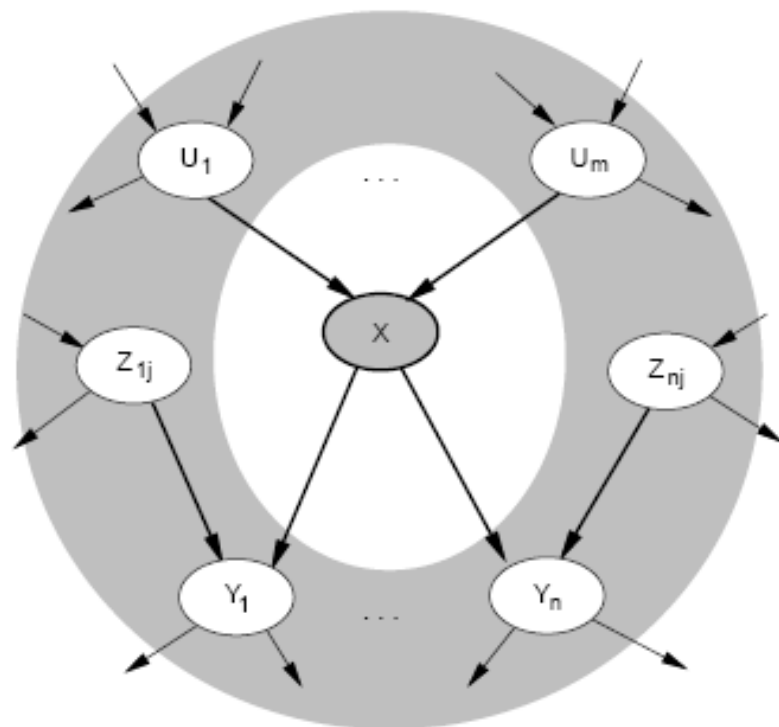
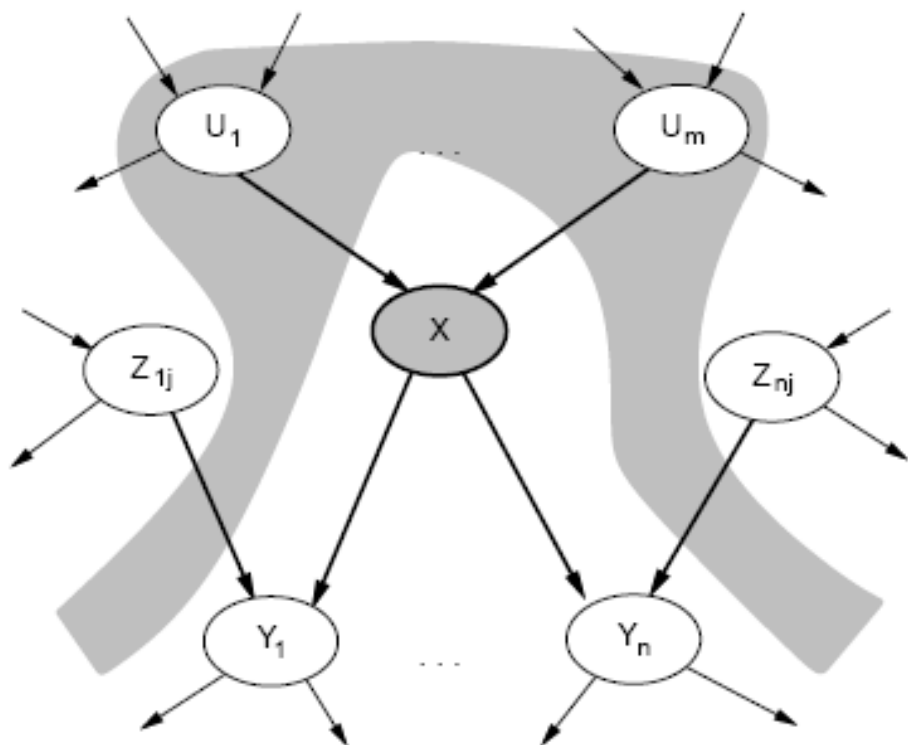
FVT-k: az együttes valószínűségi eloszlás dekomponált leírása.

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

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

Lokális (topológiai) szemantika:

- a. Minden csomópont (X) feltételesen független a nem leszármazottjaitól (Z), ha a **szülők** adottak (U).
- b. Minden csomópont (X) feltételesen független minden mástól, ha a **Markov-takarója** adott (szülők+gyerekei+gyerekeinek szülők).
- 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 háléhoz.
 - b) Legyen a $Szülő(X_i)$ 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 csomópontokhoz 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 korlátozott számú más komponenssel van kapcsolatban közvetlenül, függetlenül a komponensek teljes számától. 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 miatt lép fel, hogy a valós problémák igen strukturáltak, amit a hálók könnyűszerrel képesek kihasználni.

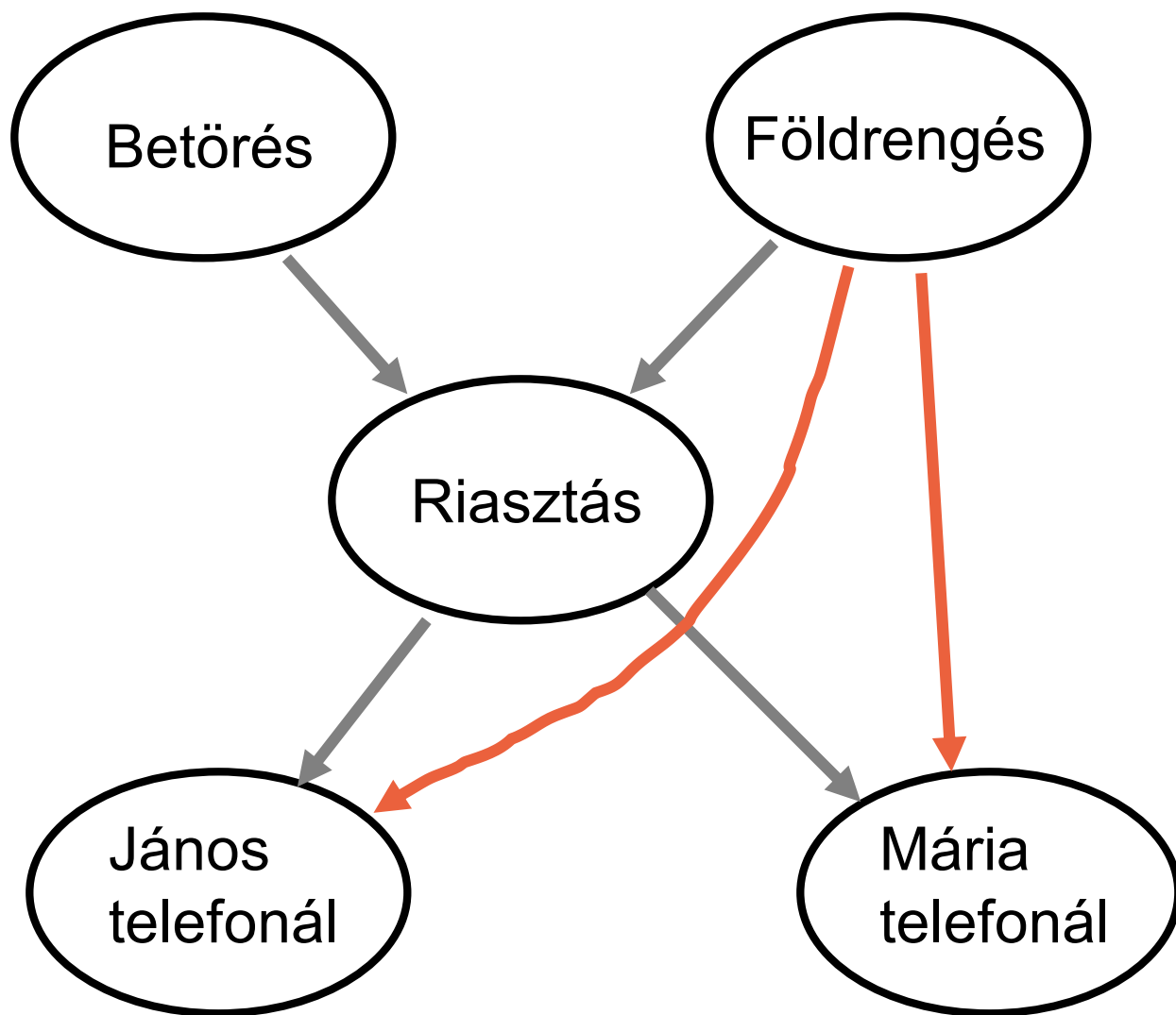
Tömörség és a csomópontok sorrendje

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.:

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



$$2 + 4 + 2 + 2 \\ = 10 \text{ (eddig)}$$

$$2 + 4 + 4 + 4 \\ = 14$$

Megéri-e?

Tömörség és a csomópontok sorrendje

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

Ezért a helyes sorrend a csomópontok hozzáadásánál: először az „alapvető okokat” adjuk a hálózathoz, 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).

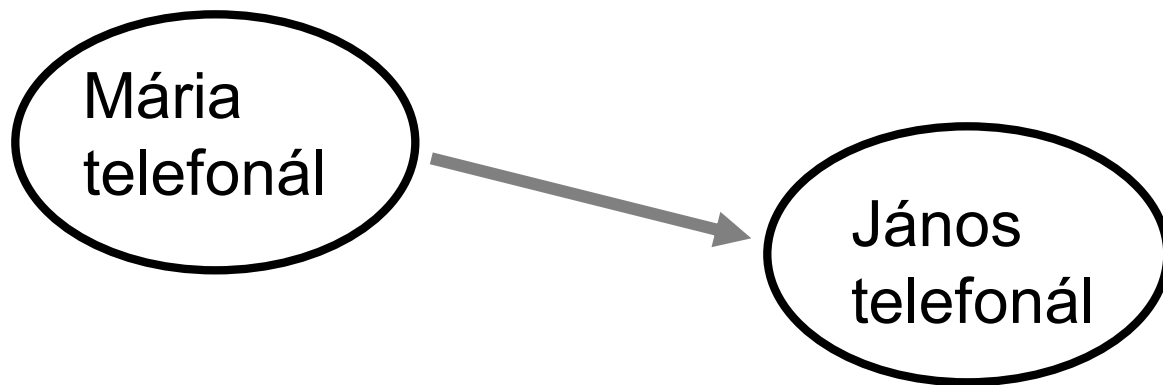
Mi történik, ha történetesen egy rossz sorrendet választunk?

pl.

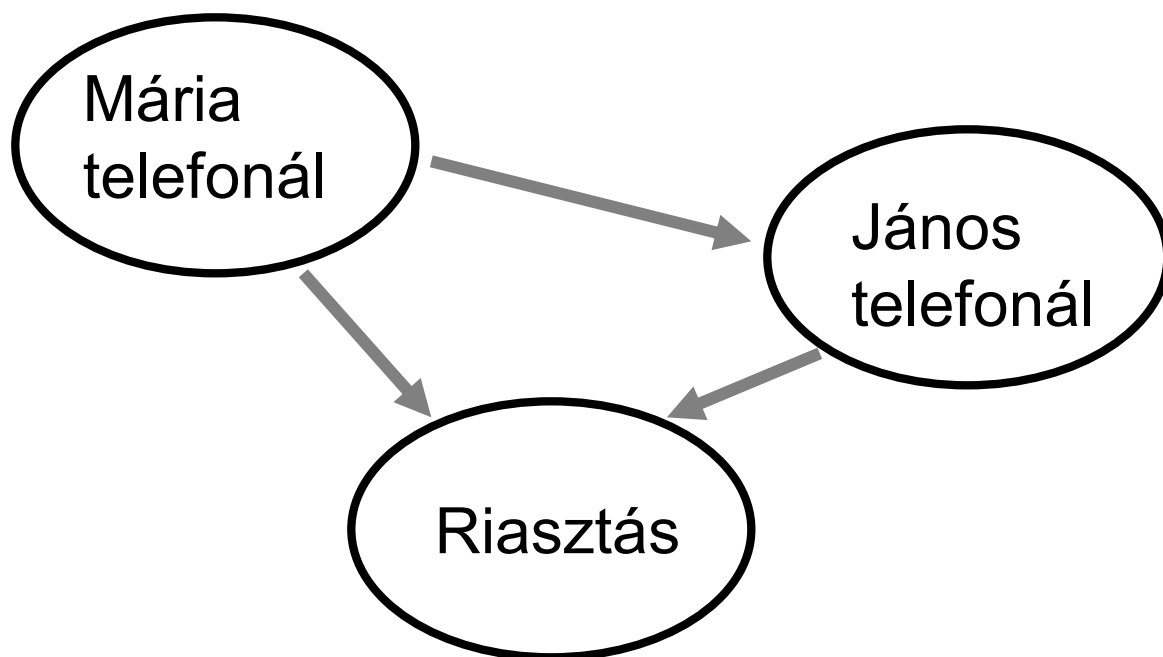
MáriaTelefonál → JánosTelefonál → Riasztás → Betörés → Földrengés.



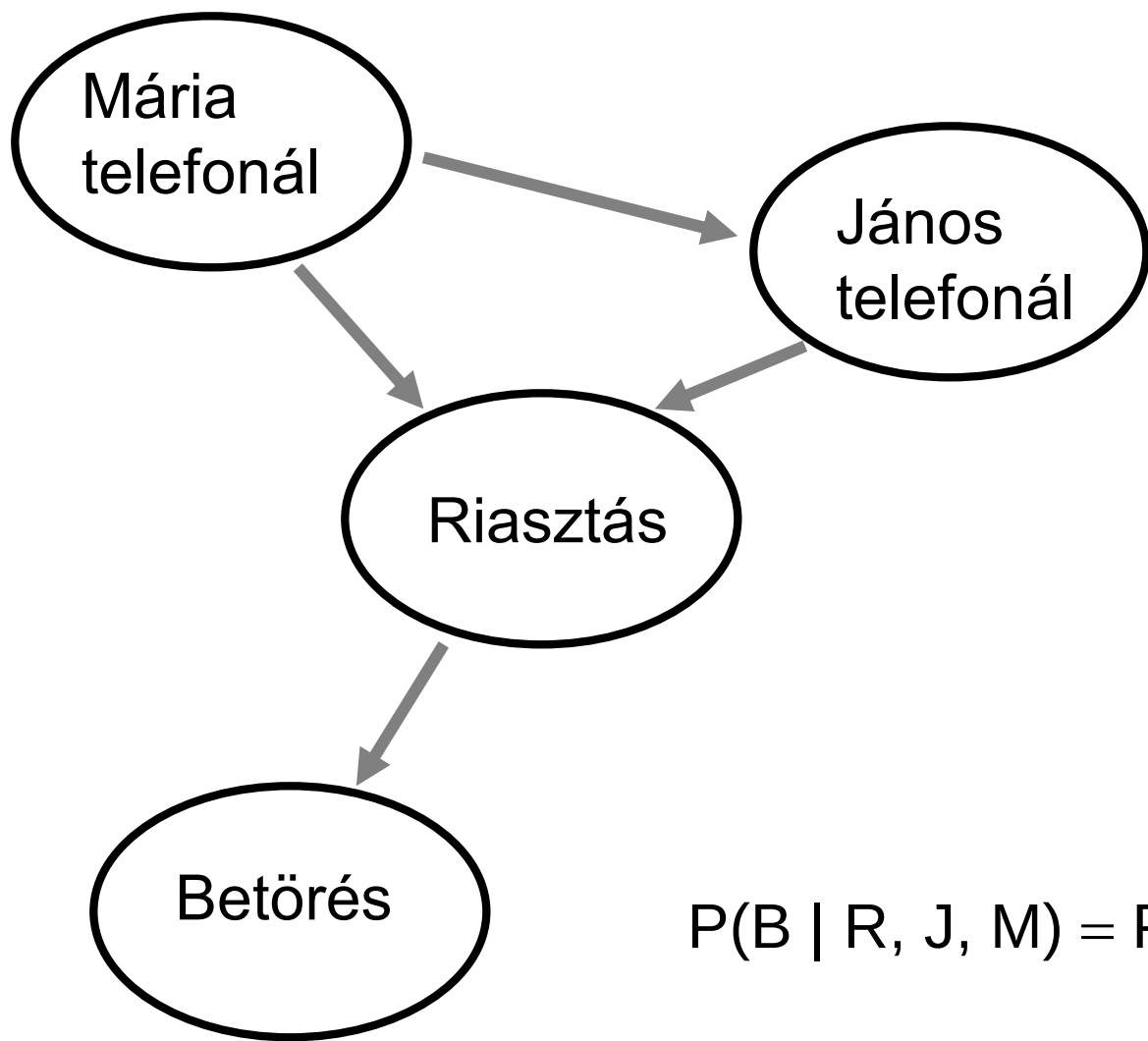
Mária
telefonál



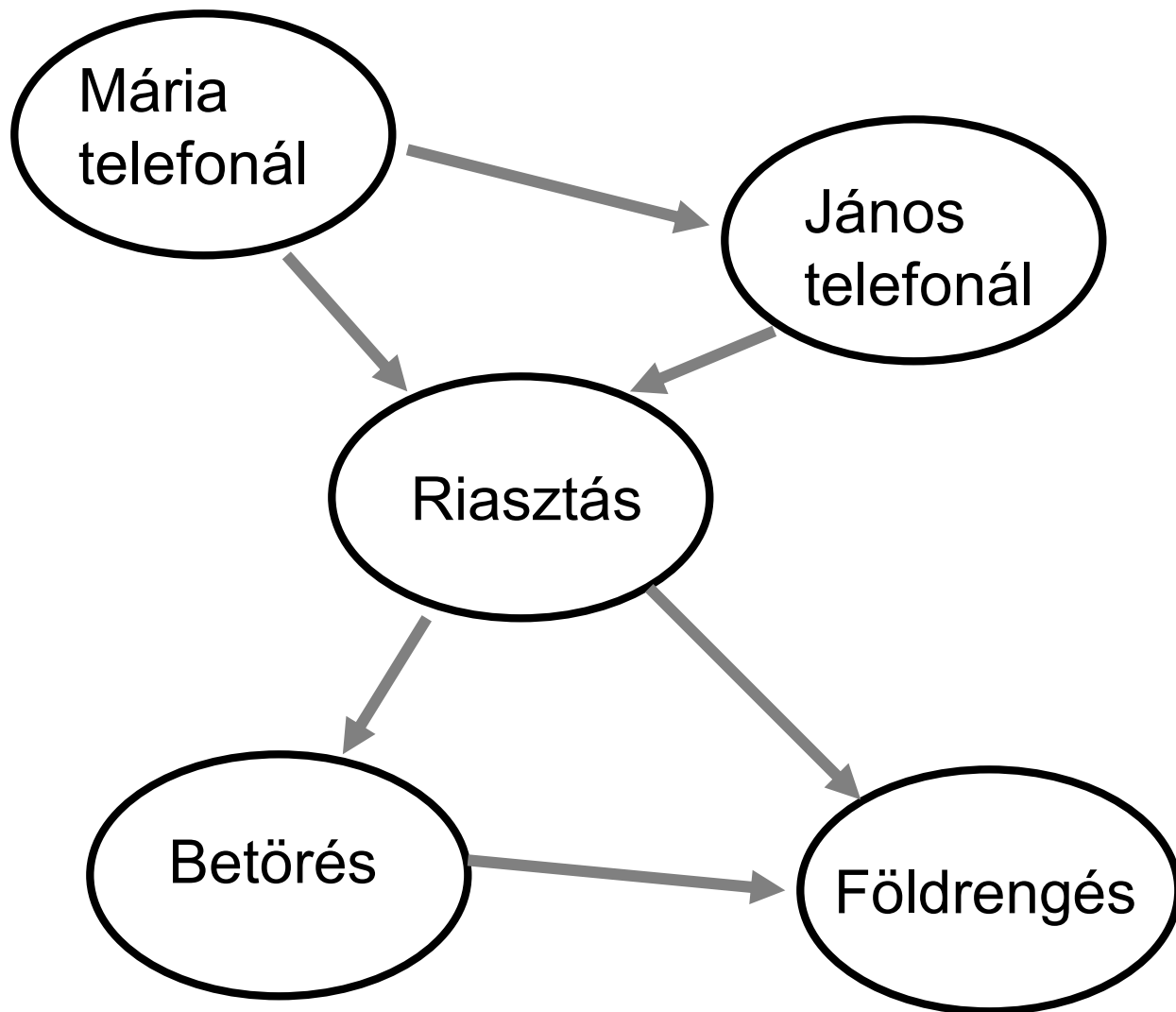
$P(J \mid M) = P(J)$? Nem (azaz van nyíl)



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



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



$$\begin{aligned} &1 + 2 + 4 \\ &+ 2 + 4 \\ &= 13 \end{aligned}$$

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

Igen Nem

Mi történik, ha történetesen egy nagyon rossz sorrendet választunk?

Mária Telefonál

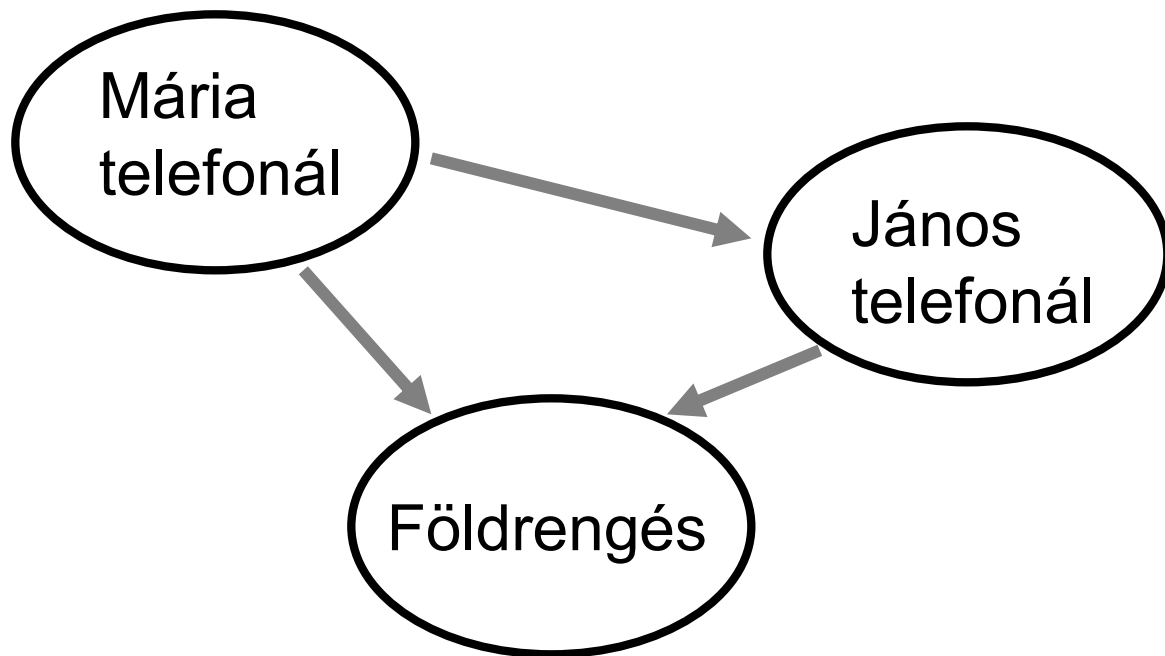
János Telefonál

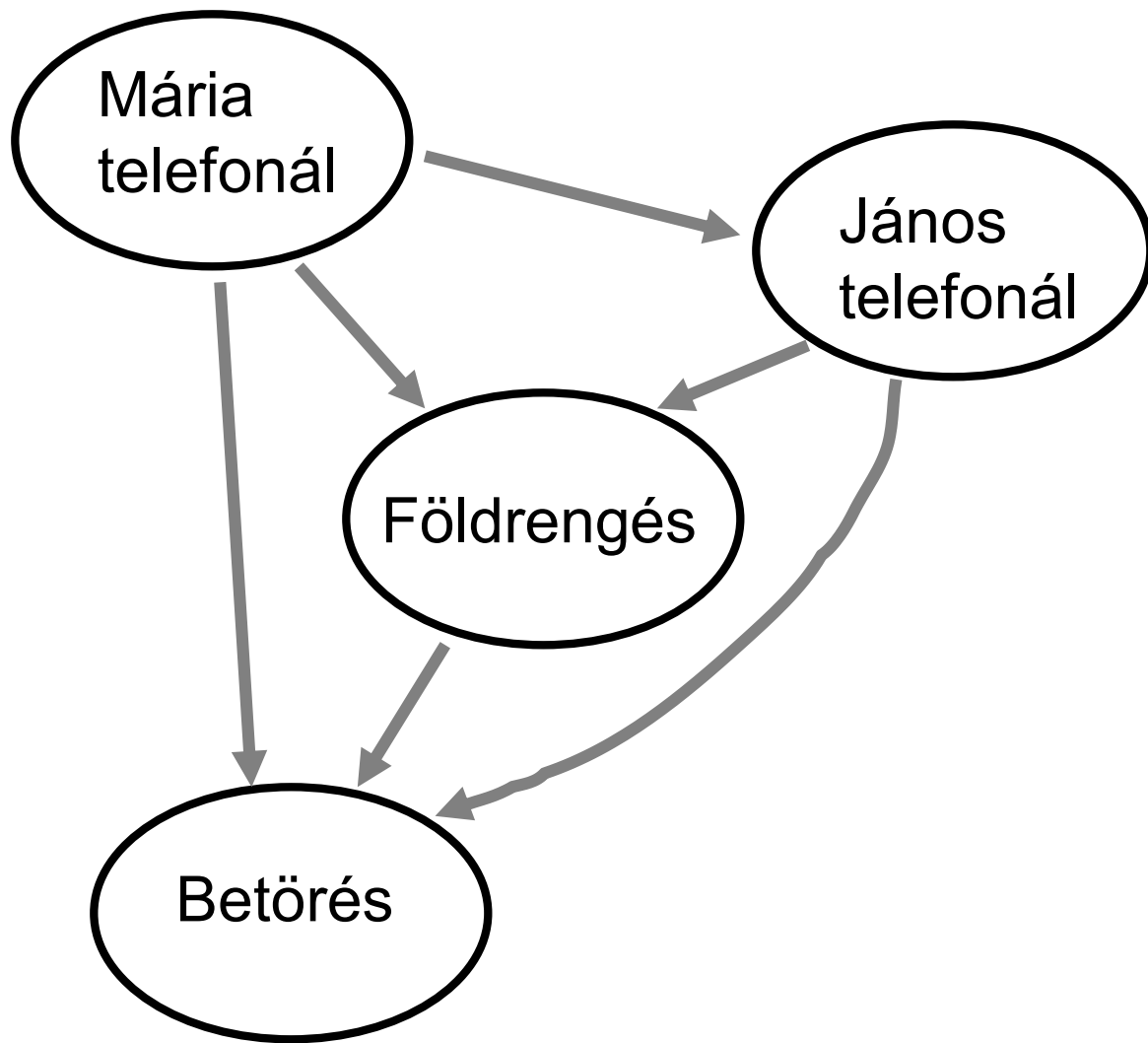
Földrengés.

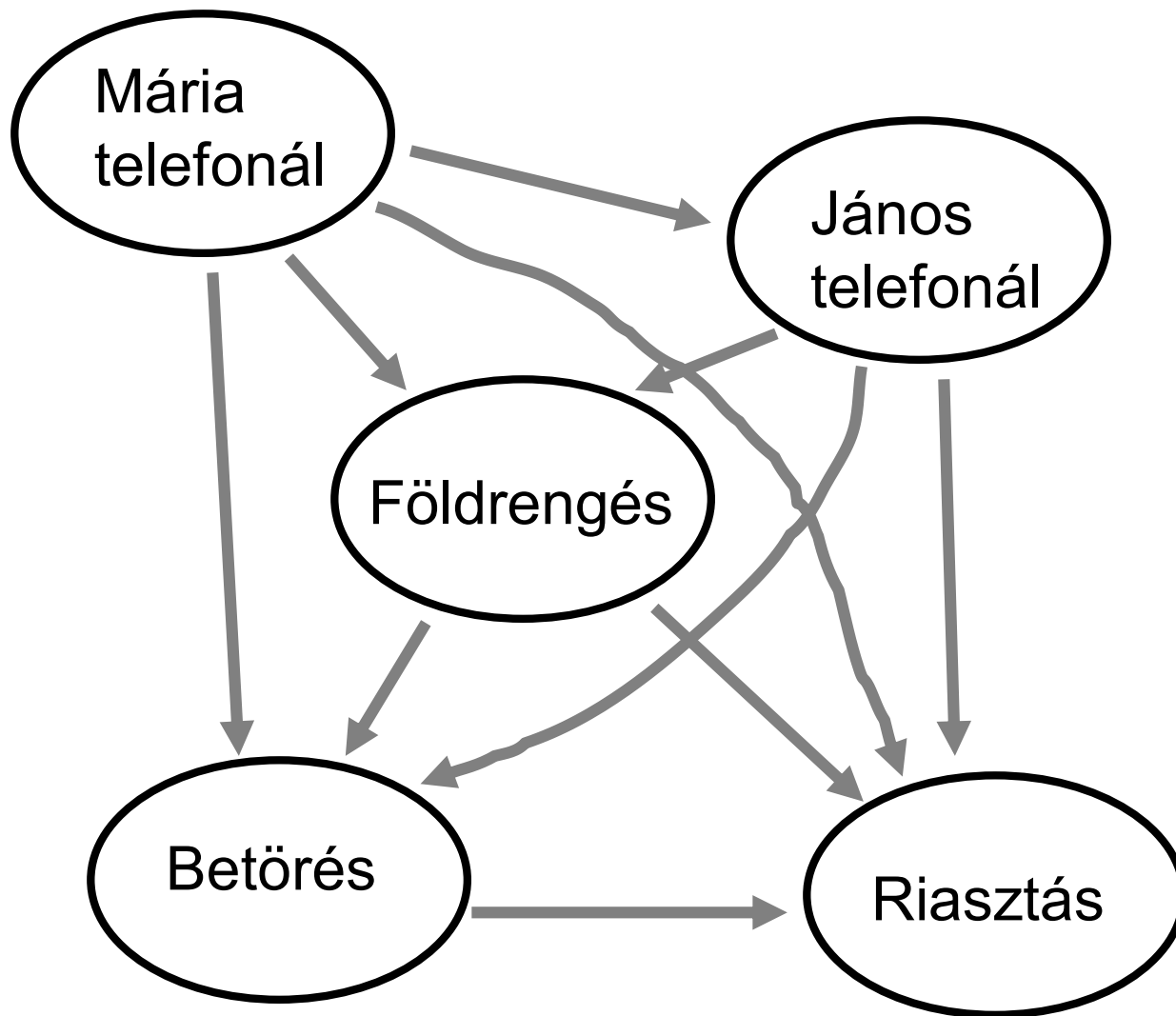
Riasztás

Betörés









$$\begin{aligned} &1 + 2 + 4 \\ &+ 8 + 16 \\ &= 31 \\ &= 2^5 - 1 \end{aligned}$$

Sok (elméletileg exponenciális számú) **valószínűség**

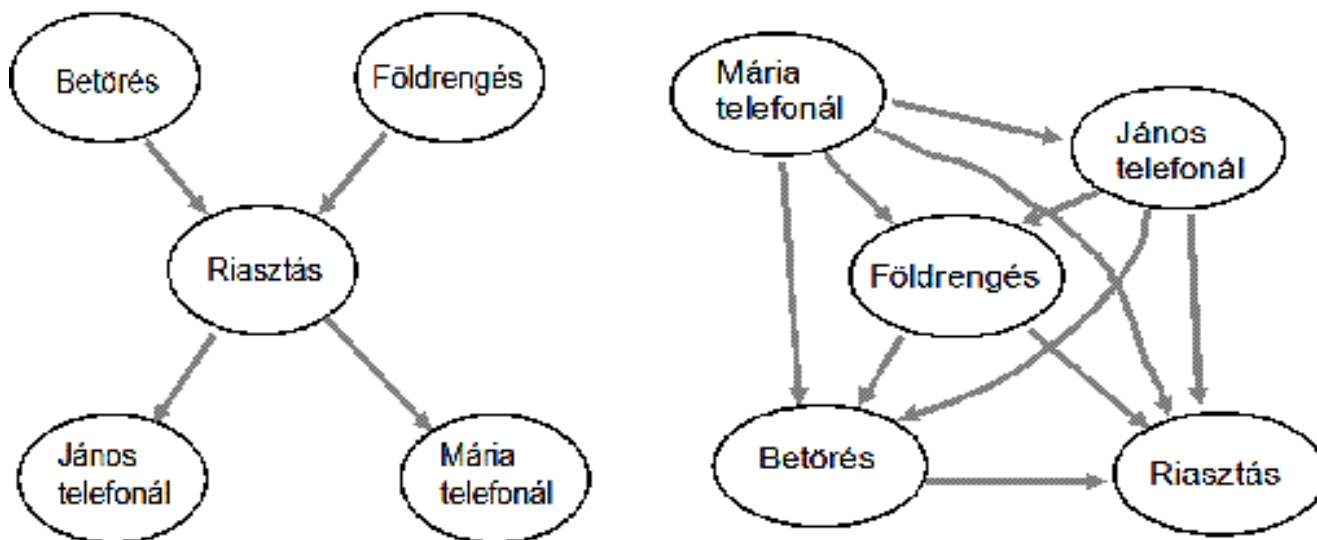
De a legrosszabb következmény:

néhány kapcsolat furcsa viszonyt reprezentál, ami nehéz és **nem természetes valószínűségi ítéleteket igényel**, pl.

$P(\text{Földrengés} \mid \text{Mária ...}, \text{János ...}) ?$

$P(\text{Betörés} \mid \text{Földrengés}) ? \dots \dots \dots$

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



Következtetés valószínűségi hálóknban

Az alapvető feladat:

kiszámítani az **a posteriori** valószínűséget a **lekérdezéses változókra**, ha a **tény** ill. **bizonyíték (evidencia)** változóknak az értékei adottak:

$$P(\text{Lekérdezéses} \mid \text{Bizonyíték})$$

A riasztós példában, pl.:

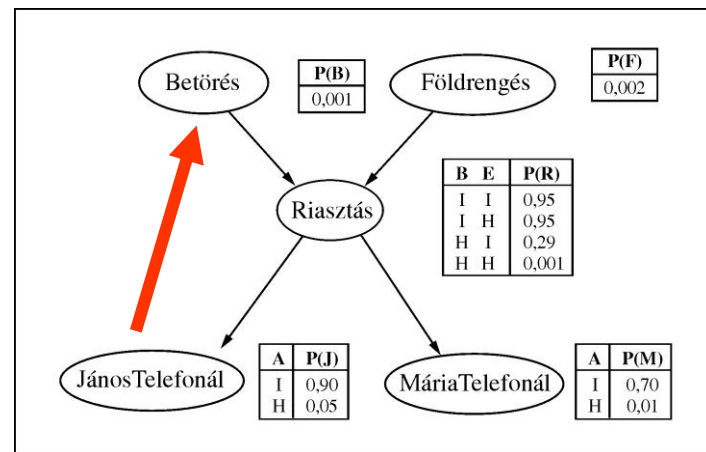
$$P(\text{Betörés} \mid \text{MáriaTelefonál és JánosTelefonál})$$

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.

Diagnosztikai következtetés (hatásról az okra)

Ha adott a *JánosTelefonál*, akkor
kiszámíthatjuk pl., hogy

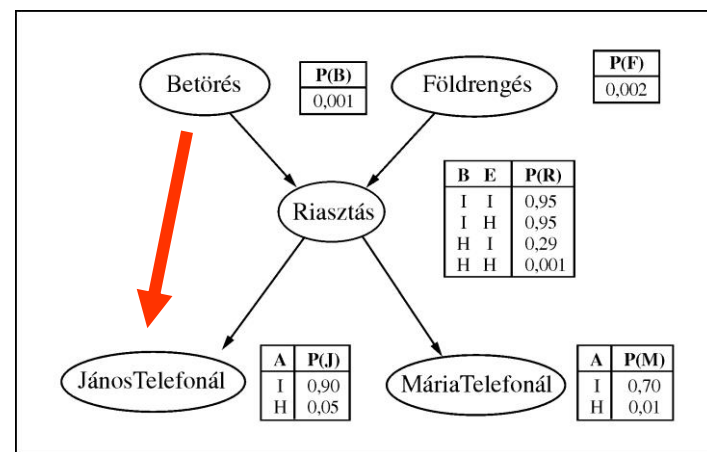
$$P(\text{Betörés} \mid \text{JánosTelefonál}) = 0,016.$$



Okozati következtetés (okról a hatásra)

Ha adott a *Betörés*, akkor
kiszámíthatjuk pl., hogy

$$P(\text{JánosTelefonál} \mid \text{Betörés}) = 0,67.$$



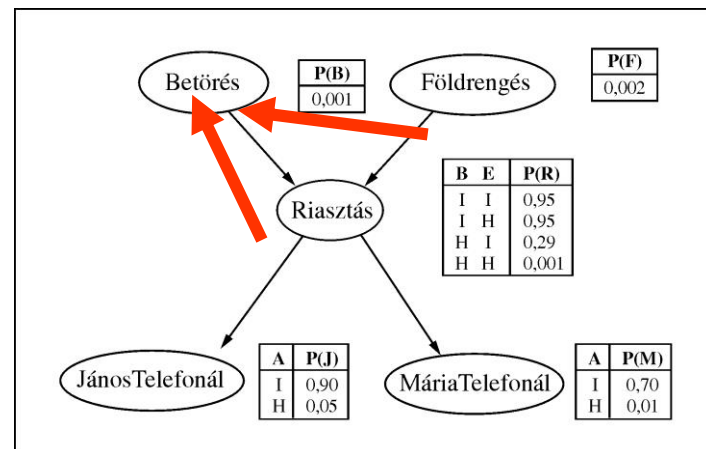
Okok közötti következtetés

(következtetés egy közös hatás okai között)

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

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

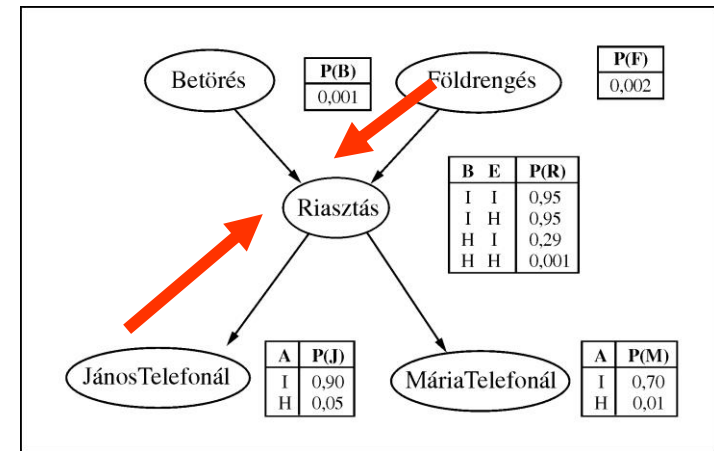
Bár a betörések és a földrengések függetlenek, az egyik jelenléte a másik valószínűségét csökkenti.
Ez a fajta következtetési mód a **kimagyarázás**.



Kevert következtetések

(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.