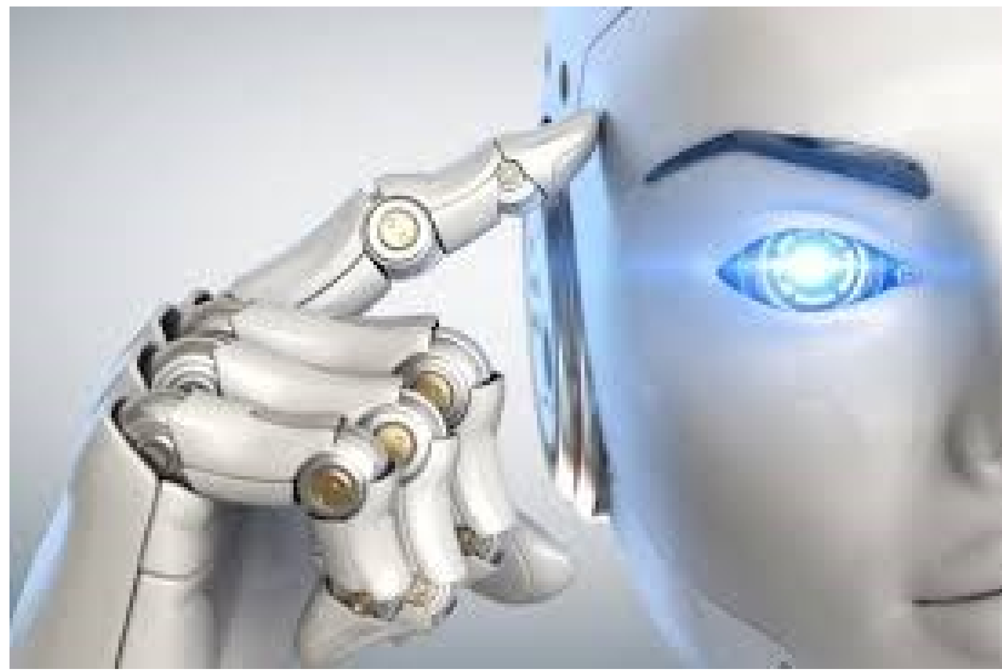
Alkalmazott mesterséges intelligencia (AMI)

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

5. ea. (2019 ősz)

Bizonytalan tudás, valószínűségszámítás

http://mialmanach.mit.bme.hu/aima/ch13
Jegyzet 13. fejezet



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

Bizonytalan tudás

Lehetséges okok

- „Lazaság/lustaság” **-** a részletes kapcsolatok megfogalmazása túl nehéz, a használatuk szintén nehézkes (véges erőforrások)
- Elméleti ismeret hiánya **-** adott problématerületnek az elméleti feltárása még nem zárult le, vagy lezárni soha nem lehet
- Gyakorlati ismeret hiánya **-** nem minden, a szabályokban hivatkozott feltétel ismert a szabályok alkalmazásakor (pl. zajos méréseket végzünk, nem ismerjük a pontos értékeket, egyes értékeket egyáltalán nem ismerünk)

Bizonytalan tudás

•Példa:

$$\forall p.\, Tünet(p, Fogfájás) \rightarrow Van(p, Fogszuvasodás) \vee$$
$$Van(p, Ínysorvadás) \vee Van(p, Bölcsességfognő) \vee \ldots$$

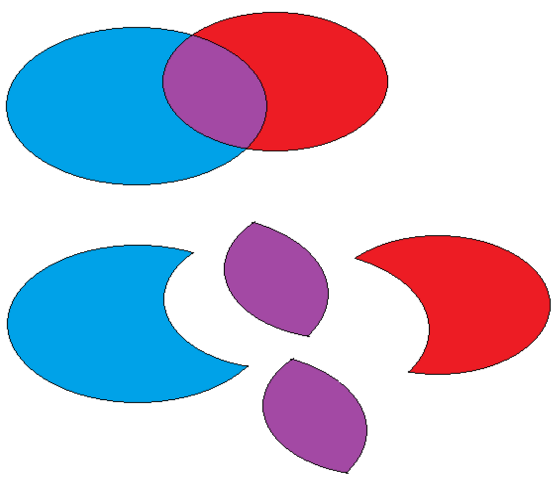
$$\exists p.\, Van(p, Fogszuvasodás) \rightarrow Tünet(p, Fogfájás)$$

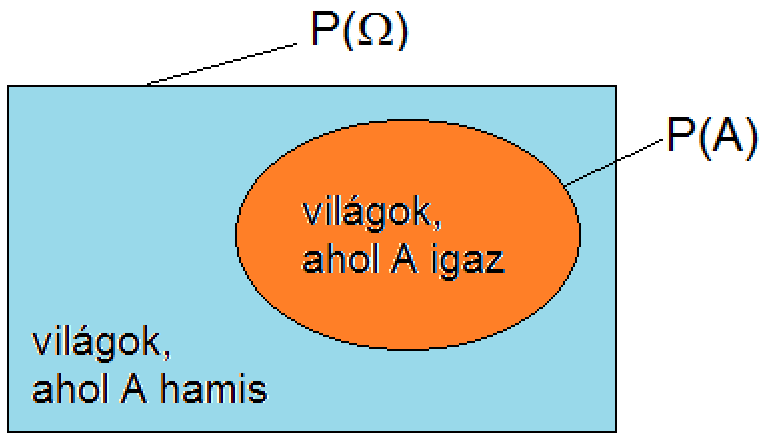
•Fogfájás és fogszuvasodás közötti kapcsolat egyik irányban sem feltétlen logikai következmény. (Pl. több különböző oka lehet a fogfájásnak.)

•Az ágens tudása legjobb esetben is csak egy bizonyos mértékű hiedelmet jelent az adott állítással kapcsolatban.

Valószínűségi axiómák

1. Minden valószínűség 0 és 1 közé esik $0 \le P(A) \le 1$.
2. A biztosan igaz állítás valószínűsége 1, a biztosan hamis állításé 0. $P(Igaz) = 1$; $P(Hamis) = 0$
3. Diszjunkció valószínűsége: $P(A \vee B) = P(A) + P(B) - P(A \wedge B)$





Valószínűségi axiómákból bizonyítható pl., hogy:

$P(\neg A) = 1 - P(A)$

$P(A) = P(A \wedge B) + P(A \wedge \neg B)$ stb.

Események – véletlen kísérlet

- **Elemi esemény:** minden lehetséges kimenetel ($e_1, e_2, \ldots, e_n$), amiről egy kísérlet elvégzése után eldönthető, hogy bekövetkezett vagy sem
  - $e_1 \wedge e_2 \wedge \ldots \wedge e_n = \emptyset$
- **Eseménytér:** egy kísérlet összes kimenetele, az összes elemi esemény halmaza (Ω).
  - $e_1 \cup e_2 \cup \ldots \cup e_n = \Omega$
- **Véletlen esemény:** az eseménytér egy részhalmaza
- **Biztos esemény:** a kísérlet során biztosan (minden kimenetelnél) bekövetkezik.
- **Ellentett esemény**: akkor és csak akkor következik be, ha az eredeti esemény nem következik be

Valószínűségi állítások

binária s	**Esemény** = Igaz	P(Esemény = Igaz) röviden P(Esemény)
többértékű (kategorikus)	**Időjárás** 1 értéket vesz fel az alábbi 4 lehetségesből {Napos, Esős, Felhős, Havazás}	P(Időjárás = Esős)
folytonos változó	**Hőmérséklet** = 22.1 ºC, **Hőmérséklet** < 22 ºC	P(Hőmérséklet < 22 ºC)

Feltételes valószínűség: P(A,B) = P(A | B) .P(B)

P(A | B) = P(A,B) / P(B)

Valószínűségi állítások átalakítása

Láncszabály: P(A,B,C,D,E) = P(A | B,C,D,E) P(B,C,D,E) =

= P(A | B,C,D,E) P(B | C,D,E) P(C,D,E) =

…

= P(A | B,C,D,E) P(B | C,D,E) P(C | D,E) P(D | E) P(E)

Teljesen általánosan:

$$P(X_1 X_2 X_3 \ldots X_N) = P(X_1 X_2 X_3 \ldots X_{N-1})\, P(X_N \mid X_1 X_2 \ldots X_{N-1})$$

$$= P(X_1 X_2 X_3 \ldots X_{N-2})\, P(X_{N-1} \mid X_1 X_2 \ldots X_{N-2})\, P(X_N \mid X_1 X_2 \ldots X_{N-1})$$

$$= \ldots = \prod_{i=2}^{N} P(X_i \mid X_1 X_2 \ldots X_{i-1}) \cdot P(X_1)$$

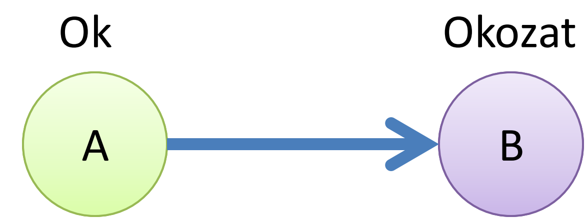
Bayes-tétel

$$P(A \mid B) = \frac{P(B \mid A) P(A)}{P(B)}$$

$$P(A_k \mid B) = \frac{P(B \mid A_k) P(A_k)}{\sum_{i=1}^{N} P(B \mid A_i) P(A_i)}, \quad \bigcup \qquad \cap A_k = \emptyset$$

Miért fontos a Bayes-tétel?

Sokszor rendelkezünk kauzális (ok-okozati) tudással:



Betegség Tünet $P(tünet \mid betegség)$

Tűz Riasztás $P(riasztás \mid tűz)$

Bayes-tétel

Viszont sokszor„ellentétes irányban” szeretnék következtetni, tehát evidenciák (tények, amik most következmények) alapján az ok(ok)ra

$P(betegség \mid tünet)$	Betegség	Tünet
$P(tűz \mid riasztás)$	Tűz	Riasztás

Bayes-tétel jelentősége

Lehetővé teszi a valószínűségi állítások átalakítását

- Így kiszámíthatóvá válnak nehezen becsülhető mennyiségek
- Oksági irány: általában könnyebb becsülni
- Diagnosztikai irány: általában nehezebb

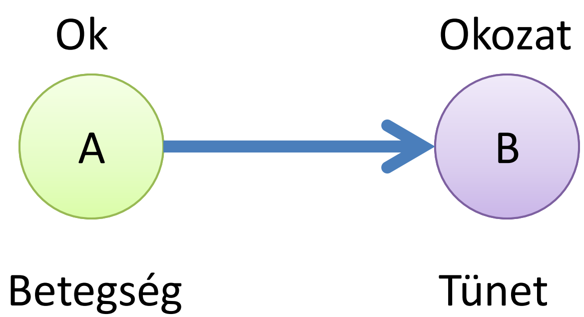
Mi a Tünet létrejöttének valószínűsége a Betegség ismeretében?

Mi a Betegség kialakulásának a valószínűsége?

$$P(Betegség \mid Tünet) = \frac{P(Tünet \mid Betegség)P(Betegség)}{P(Tünet)}$$

Mi a Betegség meglétének valószínűsége a Tünet ismeretében?

Mi a Tünet létrejöttének a valószínűsége?



Bayes-tétel jelentősége kvíz következik!

Mi a Tünet létrejöttének valószínűsége a Betegség ismeretében?

Mi a Betegség kialakulásának a valószínűsége?

$$P(Betegség \mid Tünet) = \frac{P(Tünet \mid Betegség)P(Betegség)}{P(Tünet)}$$

Mi a Betegség meglétének valószínűsége a Tünet ismeretében?

Mi a Tünet létrejöttének a valószínűsége?

a priori valószínűség (prior)

Likelihood

a posteriori valószínűség (poszterior)

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)}$$

Normalizációs konstans

(lásd később)

05.1. kvíz

Három betegség okozhatja a hajhullást: B1, B2 és B3. Mindhárom betegségben a népesség 10-10%-a szenved. Azt tapasztaltuk, hogy a B1 betegségben szenvedők 20%-nál lép fel hajhullás, a B2 betegségben szenvedők 30%-nál és a B3 betegségben szenvedők 50%-nál.
A vizsgált páciensnél hajhullást tapasztalunk. Mekkora valószínűséggel szenved az illető a B1 betegségben?

A. **0,02**

B. **0,002**

C. **2,0**

D. **0,2**

05.1. kvíz megoldás

$$P(Betegség \mid Tünet) = \frac{P(Tünet \mid Betegség)P(Betegség)}{P(Tünet)}$$

Mivel mindhárom betegség esetén azonos a P(Betegség) és a P(Tünet), ezért ugyanolyan arányt fog képviselni a három betegségre a $P(Betegség \mid Tünet)$, mint a $P(Tünet \mid Betegség)$. Tehát hajhullás esetén

$$P(B1 \mid Hajhullás) = 0,2 \quad \text{azaz} \quad (20\%)$$

Bayes-tétel – műveletek

A lehetséges értékei: $\{v_1, v_2, v_3, \ldots\ldots, v_N\}$

$$P(A = v_k) = P_k$$

$$P(A = v_k) \wedge P(A = v_j) = 0, \quad k \neq j$$

$$P(A = v_1) \vee P(A = v_2) \ldots \vee P(A = v_N) = 1$$

$$P(A = v_1) \vee P(A = v_2) \ldots \vee P(A = v_k) = \sum_{i=1}^{k} P_i$$

$$P(B) = \sum_{i=1}^{N} P(B \mid A_i) P(A_i)$$

$$\boxed{P(A \mid B) = \frac{P(B \mid A) P(A)}{P(B)} = \frac{P(B \mid A) P(A)}{\sum_{i=1}^{N} P(B \mid A_i) P(A_i)}}$$

Pl. ha A egy bináris valószínűségi változó:

$$P(A = 1 \mid B = 1) = \frac{P(B = 1 \mid A = 1) \cdot P(A = 1)}{P(B = 1 \mid A = 0) \cdot P(A = 0) + P(B = 1 \mid A = 1) \cdot P(A = 1)}$$

Bayes-tétel – műveletek **kvíz következik**

A nevező normálásra szolgál!

$$P(A_k \mid B) = \frac{P(B \mid A_k)P(A_k)}{P(B)} = \frac{P(B \mid A_k)P(A_k)}{\sum_{i=1}^{N} P(B \mid A_i)P(A_i)}$$

Ugyanis B bekövetkezése esetén (is) valamelyik A_k biztosan bekövetkezik, tehát:

$$\sum_{k=1}^{N} P(A_k \mid B) = 1 = \sum_{k=1}^{N} \frac{P(B \mid A_k)P(A_k)}{\sum_{i=1}^{N} P(B \mid A_i)P(A_i)} = \frac{\sum_{k=1}^{N} P(B \mid A_k)P(A_k)}{\sum_{i=1}^{N} P(B \mid A_i)P(A_i)}$$

Ezért gyakran nem is írjuk ki a nevezőt, csak jelezzük,hogy van egy normáló tényező (ami azt biztosítja, hogy 1 legyen az elemi események feltételes valószínűségeinek összege):

$$P(A \mid B) = \alpha \cdot P(B \mid A)P(A) \qquad \text{vagy} \qquad P(A_k \mid B) = \alpha \cdot P(B \mid A_k)P(A_k)$$

05.2. kvíz

Három betegség okozhatja a hajhullást: B1, B2 és B3. Mindhárom betegségben a népesség 10-10%-a szenved. Azt tapasztaltuk, hogy a B1 betegségben szenvedők 20%-nál lép fel hajhullás, a B2 betegségben szenvedők 30%-nál és a B3 betegségben szenvedők 50%-nál.
A népességben milyen valószínűséggel fordul elő hajhullás?

- **A.** **0,01**
- **B.** **0,15**
- **C.** **0,3**
- **D.** **0,1**

05.2. kvíz megoldás

$$P(B) = \sum_{i=1}^{N} P(B \mid A_i) P(A_i)$$

Itt most B=*Hajhullás* és A_i (i=1,2,3) a 3 betegség $B1$, $B2$, $B3$!

Mivel $P(B1) = P(B2) = P(B3) = 0{,}1$ ezért

$$P(Hajhull\acute{a}s)$$

$$= 0{,}1 \cdot \left[P(Hajhull\acute{a}s \mid B1) + P(Hajhull\acute{a}s \mid B2) + P(Hajhull\acute{a}s \mid B3) \right] =$$

$$= 0{,}1 \cdot \left[0{,}2 + 0{,}3 + 0{,}5 \right] = 0{,}1$$

Együttes valószínűség-eloszlás

• Az együttes valószínűség-eloszlás P(X_1, ... ,X_N) minden egyes elemi esemény kombinációhoz (állapothoz) valószínűséget rendel.

• Ha minden vizsgált valószínűségi változó diszkrét, akkor az együttes valószínűség-eloszlás leírható egy *N*-dimenziós táblázattal

• Egy cella = az adott állapot valószínűsége.

Írjuk le a budapesti időjárást két (egy bináris és egy ternáris) változóval:

időjárás: {napos, felhős},

hőmérséklet: {meleg, közepes, hideg}.

Akkor az együttes eloszlás P(*időjárás*, *hőmérséklet*):

Együttes valószínűség-eloszlás

Írjuk le a budapesti időjárást két (egy bináris és egy ternáris) változóval:

időjárás: {napos, felhős},

hőmérséklet: {meleg, közepes, hideg}.

Akkor az együttes eloszlás P(*időjárás*, *hőmérséklet*):

	Meleg	Közepes	Hideg
Napos	.1	.2	.1
Felhős	.05	.35	.2

- Mivel az elemi események egymást kizáróak, ezek együttes bekövetkezése szükségszerűen hamis tény.

- Az axiómákból következően: **a táblázat elemeinek összege 1.**

Marginális és más eloszlások

$$P(X,Y) = \sum_{z \in dom(Z)} P(X,Y,Z=z)$$

	Meleg	Közepes	Hideg
P(*hőmérséklet*) =	.15	.55	.3

	Napos	Felhős
P(*időjárás*) =	.4	.6

	Meleg	Közepes	Hideg
Napos	.1	.2	.1
Felhős	.05	.35	.2

Marginális és más eloszlások

P(*hőmérséklet*| *időjárás* = Napos) =

Meleg	Közepes	Hideg
.25	.50	.25

P(*időjárás*| *hőmérséklet*) =

	Meleg	Közepes	Hideg
Napos	.67	.36	.33
Felhős	.33	.64	.67

	Meleg	Közepes	Hideg
Napos	.1	.2	.1
Felhős	.05	.35	.2

Együttes valószínűség-eloszlás

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

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

$P(x_1, x_2, \ldots, x_N)$ esetén kell $2^N - 1$ független valószínűségérték, exponenciálisan növekvő számú valószínűség ismerete szükséges.

Térjünk vissza a tankönyv példájához:

$$\forall p.\, Tünet(p, Fogfájás) \rightarrow Van(p, Fogszuvasodás) \vee$$
$$Van(p, Ínysorvadás) \vee Van(p, Bölcsességfognő) \vee \ldots$$

$$\exists p.\, Van(p, Fogszuvasodás) \rightarrow Tünet(p, Fogfájás)$$

- Fogfájás és fogszuvasodás közötti kapcsolat egyik irányban sem feltétlen logikai következmény. (Pl. több különböző oka lehet a fogfájásnak.)
- Az ágens tudása legjobb esetben is csak egy bizonyos mértékű hiedelmet jelent az adott állítással kapcsolatban.
- Tegyük fel, hogy fogorvos a szondával (=hegyes fémizé) beleakad a fogunkba.

Feltételes függetlenség

Mind a fogfájásnak, mind a szonda lyukba akadásának közvetlen oka lehet a fogszuvasodás.

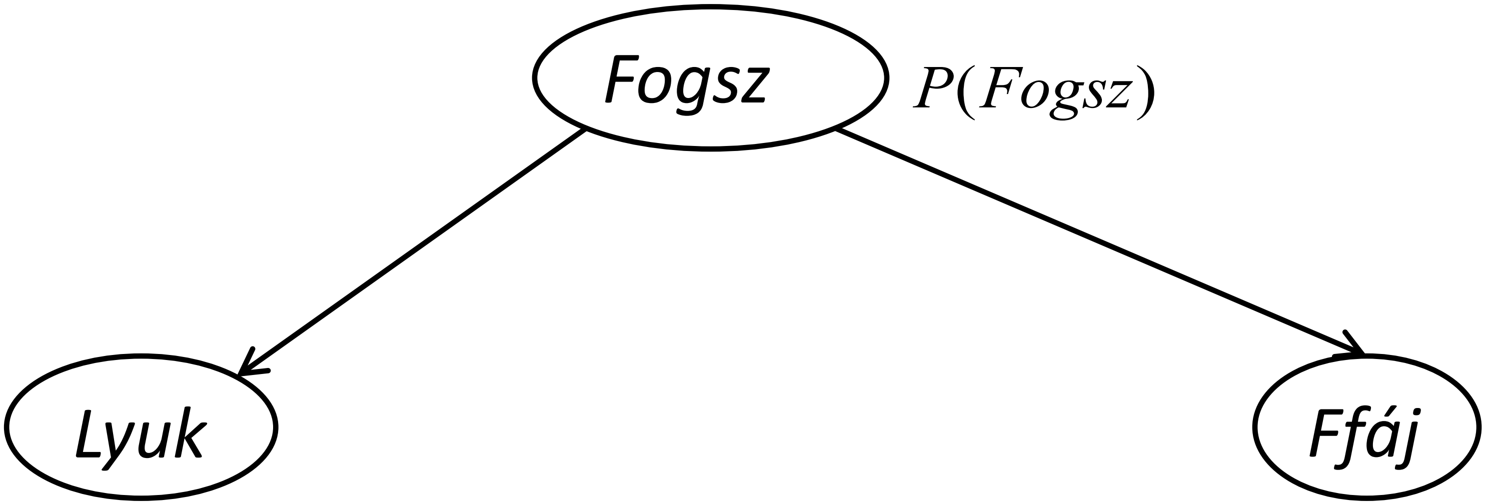
Amint **tudjuk**, hogy fogszuvasodás fennáll, nem hisszük, hogy a szonda lyukba akadásának valószínűsége a fogfájástól fog függeni.

Hasonlóképpen, a szonda találata nem befolyásolja annak valószínűségét, hogy a szuvasodás fogfájást okoz.

$$P(Lyuk \mid Ffáj \wedge Fogsz) = P(Lyuk \mid Fogsz)$$
$$P(Ffáj \mid Lyuk \wedge Fogsz) = P(Ffáj \mid Fogsz)$$

a *Fogszuvasodás* ténye esetén a *Fogfájás* és *Lyuk* között fennáll a **feltételes függetlenség**.

Feltételes függetlenség



$$P(Fogsz)$$

$$P(Lyuk \mid Fogsz) = \begin{bmatrix} P(Lyuk \mid Fogsz = 1) \\ P(Lyuk \mid Fogsz = 0) \end{bmatrix}$$

$$P(Ff\acute{a}j \mid Fogsz) = \begin{bmatrix} P(Ff\acute{a}j \mid Fogsz = 1) \\ P(Ff\acute{a}j \mid Fogsz = 0) \end{bmatrix}$$

Feltételes függetlenség - alkalmazás

$$P(Fogsz \mid Ffáj \wedge Lyuk) = P(Fogsz \mid Ffáj, Lyuk) =$$

$$= \frac{\boxed{P(Ffáj \mid Fogsz \wedge Lyuk)} \cdot P(Fogsz \wedge Lyuk)}{P(Ffáj \wedge Lyuk)} =$$

$$= \frac{\boxed{P(Ffáj \mid Fogsz)}}{P(Lyuk \mid Ffáj)} \cdot \frac{P(Lyuk \mid Fogsz) \cdot P(Fogsz)}{P(Ffáj)}$$

$$P(Lyuk \mid Ffáj \wedge Fogsz) = P(Lyuk \mid Fogsz)$$

$$P(Ffáj \mid Lyuk \wedge Fogsz) = P(Ffáj \mid Fogsz)$$

$$P(Fogsz \mid Ffáj \wedge Lyuk)$$

$$= P(Fogsz) \frac{P(Ffáj \mid Fogsz)}{P(Ffáj)} \frac{P(Lyuk \mid Fogsz)}{P(Lyuk \mid Ffáj)}$$

?

Normalizálás

Még mindig kérdéses a: $P(Lyuk|Fogfájás)$?!

- várhatóan figyelembe kell venni a tünetek összes lehetséges párosítását (hármasait, stb.), valóságban ez a kifejezés kiesik:

a nevezők szorzata: $P(Lyuk|Fogfájás)\ P(Fogfájás) = P(Lyuk)$

Ezt az un. **normalizálással** kiküszöbölhetjük, feltéve, hogy pl. a $P(Lyuk|\neg Fogszuvasodás)$-t megbecsüljük.

$$P(Fogsz \mid Lyuk) = \frac{P(Lyuk \mid Fogsz)P(Fogsz)}{P(Lyuk)}$$

$$P(Fogsz \mid Lyuk) \sim \qquad\qquad \mid Fogsz)P(Fogsz)$$

Normalizálás

$$P(Fogsz\,|\,Lyuk) = \frac{P(Lyuk\,|\,Fogsz)P(Fogsz)}{P(Lyuk)}$$

$$P(Fogsz\,|\,Lyuk) \sim \qquad\qquad |\,Fogsz)P(Fogsz)$$

$$P(Fogsz\,|\,Lyuk) = \alpha P(Lyuk\,|\,Fogsz)P(Fogsz) = C_1\alpha$$

$$P(\neg Fogsz\,|\,Lyuk) = \alpha P(Lyuk\,|\,\neg Fogsz)P(\neg Fogsz) = C_2\alpha$$

$$P(Fogsz\,|\,Lyuk) + P(\neg Fogsz\,|\,Lyuk) = (C_1 + C_2)\alpha = 1$$

$$\alpha = (C_1 + C_2)^{-1}$$

Lehetséges problémák

- Az a priori feltételes és együttes valószínűségek begyűjtése nehéz és költséges

- Az emberek rossz valószínűségbecslők

  (ha nem az események gyakoriságából próbálunk becsülni, hanem szakértőket kérünk szubjektív becslésre)

- A Bayes-szabály sok számítást igényel