V. Bizonytalanságkezelés

Bizonytalanság forrásai

- Hiányzó adat mellett történő következtetés
  - Mi lehet a páciens betegsége?
- Bizonytalan adatra épülő következtetés
  - Pontatlan műszerek pontatlan leolvasása: *80 °C ± 2 °C*
  - *objektív*
  - *szubjektív*
- Bizonytalan következtetés:
  - Lehet-e egy sárga bőrű páciens hepatitiszes?
  - Ha ismerjük *sárga bőrű hepatitiszesek / sárga bőrű betegek* arányát?
  - *objektív*
- Elmosódott jelentésű állítások:
  - Jóska magas
  - *szubjektív*
- Ellentmondó adatokból vagy ellentmondó következtetésekből származtatott következmény

Bizonytalanságkezelés alapkérdései

- Hogyan reprezentáljuk a bizonytalanságot?
  - Az ismeretekhez numerikus vagy szimbolikus értéket rendelünk
- Hogyan kombináljuk a bizonytalanságot?
  - A logikai műveletek mentén komponált összetett ismeret bizonytalanságát a komponensek bizonytalanságából számoljuk.
- Hogyan következtessünk bizonytalan információból?
  - Mennyire (milyen mértékben) bizonytalan az a következmény, amelyre bizonytalan ismeretekből következtettünk ?

Bizonytalanságkezelés technikái

- Elméleti alapokkal rendelkező modellek
  - Bayes modell
  - Valószínűségi (Bayes) háló
  - Fuzzy modell
  - Dempster-Shafer modell
- Heurisztikus módszerek
  - MYCIN
- Általánosítva a bizonytalanságkezelés fogalmát ide sorolhatók még a nemmonoton következtetések is, amelyek viszont – ellentétben a fent felsorolt numerikus technikájú módszerekkel – szimbolikus technikákat használnak.

1. Klasszikus valószínűség számítás

- A központi kérdés annak a meghatározása, hogy mi az *A* valószínűsége, ha *B* bekövetkezik.
- Feltételes valószínűség:

$$P(A \mid B) = \frac{p(A \wedge B)}{p(B)} \quad \text{ha } p(B) > 0$$

Valószínűségi változók

- Jelölések:
  - $X_i = x_i$ állításokkal dolgozunk, ahol X_i egy diszkrét valószínűségi változó, x_i pedig az értéke.
  - Speciális eset az, amikor az X_i lehetséges értékei: *igaz* vagy *hamis*. Ekkor az X_i=*igaz* helyett használhatjuk az X_i-t, az X_i=*hamis* helyett pedig $\neg X_i$ –t.

Együttes valószínűségi eloszlásfüggvény szerepe a valószínűségi következtetésben

- Feltételes valószínűségek az események együttes valószínűségi eloszlásának ismeretében kiszámolhatók.
- A gyakorlatban azonban
  - az együttes valószínűségi eloszlást nem ismerjük explicit módon,
  - ez amúgy is túl sok apriori adat tárolását igényelné, hiszen ; a memória igény exponenciálisan nő az elemi események számával,
  - ezért különféle elkerülő technikákat alkalmazunk.

Bayes tétel különféle alakjai

a) Klasszikus

$$p(B \mid A) = \frac{p(A \mid B)\, p(B)}{p(A)}$$

b) Háttértudás mellett

$$p(B \mid A,E) = \frac{p(A \mid B,E)\, p(B \mid E)}{p(A \mid E)}$$

c) Általánosított (B_1 , ..., B_n teljes és független)

$$p(B_i \mid A) = \frac{p(A \mid B_i) p(B_i)}{\Sigma_k\, p(A \mid B_k) p(B_k)}$$

Szuvas-e a fog, ha fáj és lyukas?

Russel-Norvig: AI

- $p(szuvas \mid fáj, lyukas) = ?$
  - $P(szuvas)=0.65$
  - $P(fáj \mid szuvas) =0.5$
  - $P(fáj \mid \neg szuvas)=0.1$
  - $P(lyukas \mid szuvas) =0.95$
  - $P(lyukas \mid \neg szuvas)=0.01$

Példa folytatása

- $p(szuvas \mid fáj, lyukas) = ?$

$$p(sz \mid f,ly) = \frac{p(f,ly \mid sz)\, p(sz)}{p(f,ly)}$$

- Itt a klasszikus Bayes tételt alkalmaztuk, amelyhez csak a $p(sz)$-t ismerjük. Keressünk más utat!
- Alkalmazzuk a háttér tudás melletti Bayes tételt, ahol a *fáj* tényt háttérnek tekintjük. A $p(sz \mid f)$-re pedig a közönséges Bayes tételt írjuk majd fel. (Bayes-i frissítés módszere)

$$p(sz \mid ly,f) = \frac{p(ly \mid sz,f)\, p(sz \mid f)}{p(ly \mid f)} = \frac{p(ly \mid sz,f)\, p(f \mid sz)\, p(sz)}{p(ly \mid f)\, p(f)}$$

Feltételes függetlenség

- Közönséges függetlenség
  - $p(A,B) = p(A)\, p(B)$
- Az *A* és a *B* feltételesen függetlenek a *E* fennállása esetén, ha
  - $p(A,B \mid E) = p(A \mid E)\, p(B \mid E)$
- Gyakori alkalmazási formája:
  - $p(A \mid B,E) = p(A \mid E)$
  - Ui: $p(A \mid B,E) = p(A,B,E)/p(B,E) =$
    $= p(A,B \mid E)\, p(E)/p(B \mid E)\, p(E) = p(A,B \mid E)/p(B \mid E) =$
    $= p(A \mid E)\, p(B \mid E)/p(B \mid E) = p(A \mid E)$

Példa folytatása

- Feltételezhetjük azt, hogy a lyukas fog a szuvasodás következménye, és nem függ a fogfájástól, azaz a *ly* feltételesen független az *f*-től az *sz*-re nézve

$$p(ly \mid sz,f) = p(ly \mid sz)$$

- Összeolvasva

$$p(sz \mid ly,f) = \frac{p(ly \mid sz)\, p(f \mid sz)\, p(sz)}{p(ly \mid f)\, p(f)}$$

- Itt csak a nevezőbeli valószínűségeket nem ismerjük, viszont a számlálóbeli sz feltételre felírt valószínűségek mellett ismertek a $\neg sz$ feltétel melletti valószínűségek Ilyenkor alkalmazhatjuk a normalizálás technikáját .

Példa folytatása

$$p(sz \mid ly,f) = \frac{p(ly \mid sz)\, p(f \mid sz)\, p(sz)}{p(f)\, p(ly \mid f)} = \alpha\, p(ly \mid sz)\, p(f \mid sz)\, p(sz)$$

$$p(\neg sz \mid ly,f) = \frac{p(ly \mid \neg sz)\, p(f \mid \neg sz)\, p(\neg sz)}{p(f)\, p(ly \mid f)} = \alpha\, p(ly \mid \neg sz)\, p(f \mid \neg sz)\, p(\neg sz)$$

$$1 = p(sz \mid ly,f) + p(\neg sz \mid ly,f) = \alpha\,[\, p(ly \mid sz)\, p(f \mid sz)\, p(sz) + p(ly \mid \neg sz)\, p(f \mid \neg sz)\, p(\neg sz)]$$

$$\alpha = 1/[p(ly \mid sz)\, p(f \mid sz)\, p(sz) + p(ly \mid \neg sz)\, p(f \mid \neg sz)\, p(\neg sz)]$$

Példa befejezése

- $P(lyukas \mid szuvas) = 0.7$ $\quad P(lyukas \mid \neg szuvas) = 0.01$
- $P(fáj \mid szuvas) = 0.5$ $\quad P(fáj \mid \neg szuvas) = 0.1$
- $P(szuvas) = 0.65$

$$\alpha = 1/[p(ly \mid sz)\, p(f \mid sz)\, p(sz) + p(ly \mid \neg sz)\, p(f \mid \neg sz)\, p(\neg sz)]$$
$$= 1/[\, 0.7*0.5*0.65 + 0.01*0.1*0.35\,] = 0.0001225$$

$$p(sz \mid ly,f) = \alpha\, p(ly \mid sz)\, p(f \mid sz)\, p(sz) = \alpha\, 0.7*0.5*0.65$$

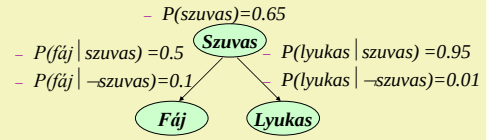
$$p(sz \mid ly,f) = \frac{0.2275}{0.227535} = 0.99846$$

Bayes modell értékelése

- Még a bevetett trükkök ellenére is sok a priori valószínűséget kell tárolni.
- Az a priori valószínűségekhez nehéz hozzájutni.
- A következtetés túl ötletszerűnek tűnik, ezért nehéz algoritmizálni.
- Matematikailag jól megalapozott, de igen számításigényes, és magyarázatadásra nem alkalmas.
- A modell új ismeretekkel nehezen bővíthető. Nem elég az új esemény és a vele kapcsolatos feltételes események valószínűségeit megadni, hanem a korábbi valószínűségi értékeket is felül kell bírálni.

2. Bayes (valószínűségi) hálók

- Az előző példa megoldásánál alkalmazott módszer általánosítható lenne, ha a feltételes függetlenségek felismerése nem lenne esetleges.
- Egy tömör reprezentáció kellene, amelyből könnyű kiolvasni a feltételes függetlenségeket.



Reprezentáció Bayes hálóval

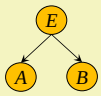
- Tekintsük a tárgyprobléma valószínűségi változóit.
- Feleltessük meg a változókat egy körmentes irányított gráf csúcsainak.
- Ábrázoljuk az irányított élekkel a változók közötti közvetlen ok-okozati összefüggéseket (ez által feltárjuk a feltételes függetlenségeket).
- Adjuk meg az csúcsok feltételes valószínűségi tábláit (FVT): $p(X_i = x_i \mid szülő(X_i) = x_{i1}, \ldots, x_{ik})$
  ahol a $szülő(X_i)$ az X_i változó csúcsának szülőcsúcsaihoz rendelt $X_{i1}, \ldots, X_{ik}$ változók együttese.

Bayes háló kifejező ereje

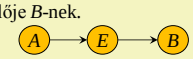
- Az együttes valószínűségi eloszlás (a lánc-szabály alapján)
  $p(X_1 = x_1, \ldots, X_n = x_n) =$
  $= P(X_n = x_n \mid X_1 = x_1, \ldots, X_{n-1} = x_{n-1}) * p(X_1 = x_1, \ldots, X_{n-1} = x_{n-1}) =$
  $= P(X_n = x_n \mid X_1 = x_1, \ldots, X_{n-1} = x_{n-1}) *$
  $*P(X_{n-1} = x_{n-1} \mid X_1 = x_1, \ldots, X_{n-2} = x_{n-2}) * \ldots * p(X_2 \mid X_1) * p(X_1)$
- Sorszámozzuk meg úgy a változókat, hogy ha $i>j$, akkor X_i-ből ne vezessen irányított út X_j-be, azaz minden i-re $szülő(X_i) \subseteq \{X_1, \ldots, X_{i-1}\}$.
- Ekkor a feltételes függetlenség miatt
  - $p(X_i = x_i \mid X_1 = x_1, \ldots, X_{i-1} = x_{i-1}) = p(X_i \mid szülő(X_i) = x_{i1}, \ldots, x_{ik})$
- tehát az adott tárgykör együttes valószínűségi eloszlása a Bayes háló FVT-iből közvetlenül megkapható.
  - $p(X_1 = x_1, \ldots, X_n = x_n) = \Pi_{i=1\ldots n}\, p(X_i = x_i \mid szülő(X_i) = x_{i1}, \ldots, x_{ik})$

Feltételes függetlenség speciális esetei Bayes hálóban

- Feltételes függetlenség: $p(A,B \mid E) = p(A \mid E)\, p(B \mid E)$
  - Ha E szülője A-nak is és B-nek is.



  - Ha A szülője E-nek, és E szülője B-nek.

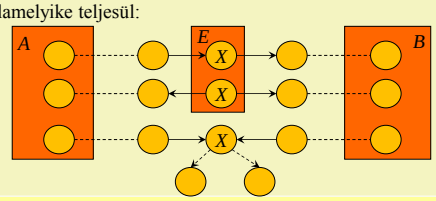


  - Ha A szülője E-nek, és E szülője B-nek



Feltételes függetlenség felismerése Bayes hálóban

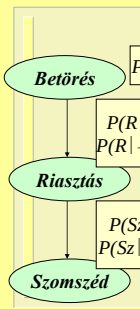
- Legyenek A, B és E összetett események, azaz mindegyikhez több csúcs tartozhat a Bayes hálóban.
- Az A és B feltételesen független az E-re nézve, ha minden A és B-beli csúcs közti irányítatlan útvonalra az alábbi esetek valamelyike teljesül:



Bayes hálók tervezése

- Határozzuk meg a tárgytartományt leíró változók halmazát, majd meghatározott sorrendben dolgozzuk fel őket:
  1. Válasszunk ki olyat, amely kizárólag a hálóhoz csatolt változóktól függ, és új csúcsként vegyük fel a hálóba
  2. A hálóbeli változóknak vegyük azt a minimális halmazát, amelyek közvetlenül hatnak az új változóra. Rajzoljuk be ezeket a függőségeket reprezentáló éleket.
  3. Töltsük ki az új csúcs FVT-jét.

A szomszédunk telefonált, hogy szól a betörés-riasztónk a lakásunkban. Betörtek volna hozzánk? Russel-Norvig: AI



$p(B \mid Sz) = p(Sz \mid B)p(B) / p(Sz)$ — Bayes tétel

$= \alpha\, p(Sz \mid B)p(B)$ — normalizálás

$= \alpha\, [p(Sz,R \mid B) + p(Sz,\neg R \mid B)]\, p(B)$ — telj fgl rsz

$= \alpha\, [p(Sz \mid R,B)\, p(R \mid B) + p(Sz \mid \neg R,B)\, p(\neg R \mid B)]\, p(B)$ — lánc szabály

$= \alpha\, [p(Sz \mid R)\, p(R \mid B) + p(Sz \mid \neg R)\, p(\neg R \mid B)]\, p(B)$ — felt. fgl.

$= \alpha\, 0.0008575$

$p(\neg B \mid Sz) = \alpha\, 0.0507991$

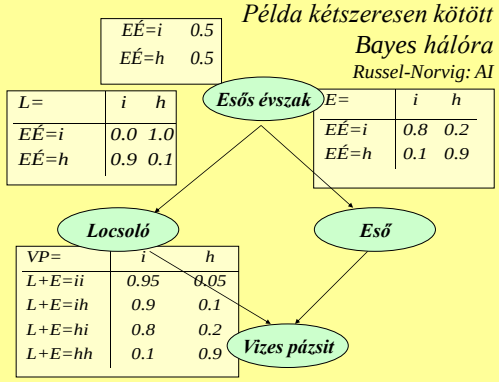
$\alpha = 19.3585$ — normalizálás vége

$p(B \mid Sz) = 0.0166$

Következtetés Bayes hálókban

- Célja egy feltételes valószínűség meghatározása a Bayes módszerre alapuló számítással (Bayes tételek, normalizálás, felbontás teljes fgl. eseményrendszerre, lánc-szabály, feltételes fgl)
- Egy feltételes valószínűség kiszámolására egy (rekurzív) algoritmus készíthető, amelynek számításigénye erősen függ a háló bonyolultságától.
- Egyszeresen kötött hálókra (fa-gráfokra), ahol két csúcs között nincsenek alternatív irányítatlan útvonalak, van lineáris futási idejű algoritmus.
- Többszörösen kötött hálók esetén különféle redukáló módszereket alkalmazhatunk.

Példa kétszeresen kötött Bayes hálóra Russel-Norvig: AI



Következtetés többszörösen kötött hálókban

- Összevonásos eljárások
  - változók összevonásával fa-gráfot készítünk.
- Vágóhalmaz feltételezésen alapuló eljárások
  - a hálóból több fa-gráfot készítünk úgy, hogy bizonyos változókat elhagyva, azok egy-egy értékét rögzítve készítünk fa-gráfot, és a válasz az így kapott fa-gráfokból nyert eredmények (elhagyott változók adott értékre vonatkozó valószínűségével) súlyozott átlaga lesz.
- Sztochasztikus szimulációs eljárások
  - A háló valószínűségi értékekeit figyelembe véve példákat generálunk. A válasz a jó példák relatív gyakorisága lesz.

$$p(A \mid B) = \frac{A \text{ és } B\text{-nek megfelelő példák száma}}{B\text{-nek megfelelő példák száma}}$$

Bayes hálók tanulása

- FVT tanulása ismert háló-struktúra mellett példák (azaz változók által felvett érték-együttesek) alapján:
  - Az egyes feltételes valószínűségeket a példák száma alapján mért relatív gyakorisági értékkel becsüljük.
    - Probléma: ha a példák hiányosak, azaz nem ismerjük, hogy egy példában bizonyos változó milyen értéket vesz fel.
- Bayes háló szerkezetének tanulása:
  - Metrikát definiálunk a feladat és az azt leíró háló „távolságára" és ez alapján keressük a legjobban illeszkedő struktúrát.

Bayes hálók értékelése

- Kevesebb a priori valószínűséget kell benne tárolni, mintha az együttes valószínűségi eloszlásfüggvényt akarnánk ábrázolni.
- Egyszerűen bővíthető anélkül, hogy eddigi valószínűségeket újra kellene gondolni.
- A következtetés felhasználható magyarázatadásra.
- Matematikailag jól megalapozott, de - az erőfeszítéseink ellenére - igen számításigényes.

3. Heurisztikus technikák

- „Betörés-riasztó-szomszéd" probléma:
  - szabályok:
    ha *a szomszéd hallja a riasztót* ***akkor*** *szól a riasztónk*
    $Sz \rightarrow R$ (0.95)
    ha *szól a riasztónk* ***akkor*** *betörtek hozzánk*
    $R \rightarrow B$ (0.999)
  - tény: *a szomszéd telefonál, hogy hallja a riasztót*
    Sz
- Betörtek-e hozzánk?
  - Sz , $Sz \rightarrow R \Rightarrow R$; R , $R \rightarrow B \Rightarrow B$
  - $T(p)$, $T \rightarrow K(q) \Rightarrow K(p*q)$
  - $Sz(1) \Rightarrow R(0.95) \Rightarrow B(0.95*0.999)$

MYCIN bizonyossági mértéke

- A klasszikus valószínűségi értékek helyett közelítő értéket - bizonyossági faktort - rendelünk az egyes állításokhoz.
  - h bizonyossági faktora $CF[h] = MB[h]-MD[h]$
  - $MB[h]$ ~ a h-ban való hit mértéke
  - $MD[h]$ ~ a h-ban való kételkedés mértéke
- Az $e \rightarrow h$ alakú állítások mértékei
  - e fennállása esetén a h bizonyossági faktora: $CF[h,e]=MB[h,e]-MD[h,e]$

MYCIN kalkulusa

- Következmény mértéke
  - $MB[h] = MB[e] * MB[h,e]$
- Összetett állítások mértékei
  - $MB[h_1 \wedge h_2,e] = min(MB[h_1,e],MB[h_2,e])$
  - $MB[h_1 \vee h_2,e] = max(MB[h_1,e],MB[h_2,e])$
- Interferencia szabály
  - $MB[h,e_1 \wedge e_2] = \begin{cases} MB[h,e_1] + MB[h,e_2](1 - MB[h,e_1]) \\ 1 \quad \text{ha } MD[h,e_1 \wedge e_2] = 0 \end{cases}$
  - $MD[h,e_1 \wedge e_2] = \begin{cases} MD[h,e_1] + MD[h,e_2](1 - MD[h,e_1]) \\ 0 \quad \text{ha } MB[h,e_1 \wedge e_2] = 1 \end{cases}$

MYCIN modell értékelése

- Előnyei:
  - Közel áll az ember valóságszemléletéhez. Ezt sokszor azzal is erősítik, hogy a [0,1] intervallum helyett [0,100] skálával dolgoznak.
  - Könnyű vele számolni.
- Hátrányai:
  - Nem általánosítható a használata.
  - Nem rendelkezik jól megalapozott elmélettel. Egy attribútum lehetséges értékeinek CF-jeit összeadva nem 1 jön ki eredményül.

4. Fuzzy következtetés

- Tudjuk, hogy
  - HA tananyag nehézsége=közepes ÉS
    jegyzeteltsége=rossz
    AKKOR tanulási_idő=hosszú
  - HA tananyag nehézsége=könnyű ÉS
    jegyzeteltsége=rossz
    AKKOR tanulási_idő=közepes
- Mennyi ideig tanuljuk a *45%*-ban nehéz, *31%*-osan jegyzetelt tananyagot?

Fuzzy halmaz

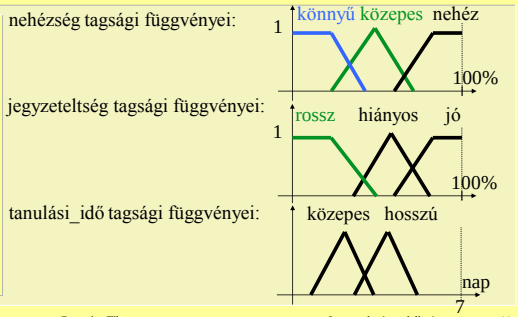
- Azt eldönteni, hogy egy elem egy adott kategóriába besorolható-e nem mindig egyértelmű
  - Egy pont félig nehéz tananyag biztosan (100 %-osan) mondható közepes nehézségűnek. De mit mondhatunk egy harmad nehézségű tananyagról? Ez mennyire (hány százalékosan) közepes nehézségű?
- Egy fuzzy (gyengén definiált) halmazhoz tartozó elemekhez hozzárendelünk egy 0 és 1 közötti értéket annak kifejezésére, hogy mennyire tekinthető biztosnak az, hogy az elem a halmazban van.
  - közepes nehézségű tananyag fogalma tehát egy fuzzy halmaz, ahol egy konkrét tananyag ennek egy eleme.

Tagsági függvény

- Az *A* fuzzy halmazhoz tartozó tagsági függvény egy $\mu_A: X \to [0,1]$ leképezés.
  - Megadja egy elemről (pl. a 33%-osan nehéz tananyagról), hogy milyen mértékben tartozik egy fuzzy halmazhoz (pl. a közepesen nehéz tananyagokhoz).
- $A = \{(x, \mu_A(x)) \mid x \in X\}$



A példa tagsági függvényei



Fuzzy állítás

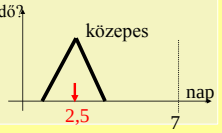
- Minden fuzzy halmazhoz megadható az a fuzzy állítás, amely azt dönti el, hogy egy elem a fuzzy halmazhoz tartozik-e. Ennek kimenetele nem az igaz és hamis értékek, hanem a *[0,1]* intervallum egy eleme.
  - Egy 50%-os nehézségű tananyag biztosan (*100%*), azaz *1* mértékben mondható közepesnek,
  - egy 16,5%-os nehézségű tananyag már egyáltalán nem közepes nehézségű, azaz 0 mértékben tartozik oda,
  - a 33%-os nehézségű tananyagról csak ½ mértékben (50%) mondhatjuk, hogy közepes.

Terminológia

- Egy fuzzy állítás egy elemről (pl. egy 33%-os nehézségű tananyagról) dönti el, hogy az egy fuzzy halmazhoz (pl. közepesen nehéz tananyagok közé) tartozik-e.
- Egy fuzzy halmazt egy nyelvi változó és annak értéke, a nyelvi kifejezés együttesen határozza meg.
  - nyelvi változó ~ pl. tananyag nehézsége
  - nyelvi kifejezés ~ pl. könnyű, közepes, nehéz.
- Finomítás céljából nyelvi kifejezésekre nyelvi módosítókat (nagyon, többé-kevésbé, stb.) is alkalmazhatunk
  - Pl: $\mu_{nagyon\ könnyű}(x) = \mu_{könnyű}(x)^2$,
    $\mu_{többé_kevésbé\ nehéz}(x) = \mu_{nehéz}(x)^{½}$

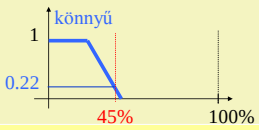
Fuzzifikálás, defuzzifikálás

- Fuzzifikálás az folyamat, amikor kiszámítjuk, hogy egy elem melyik fuzzy halmazhoz milyen mértékben tartozik.
  - Pl: egy 33%-os nehézségű tananyag milyen mértékben tartozik a közepes, milyen mértékben a könnyű tananyagok közé.
- Defuzzifikálás során egy fuzzy halmaznak keressük a legjellemzőbb elemét, amelynek a fuzzy halmazhoz tartozásának mértéke a legnagyobb.
  - Pl: Hány nap a közepes tanulási_idő?
  - Súlypont (centroid) módszer



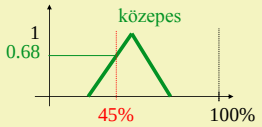
A példa fuzzifikálása

- Kérdés: Mennyi ideig tanuljuk a *45%*-ban nehéz, *31%*-osan jegyzetelt tananyagot?
- Először eldöntjük, hogy az adott fuzzy halmazok közül melyikhez,milyen mértékben tartozik a tananyag?
  - könnyű nehézségű (KN) *0.22* mértékben

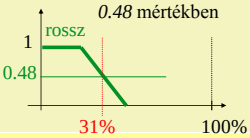


A példa fuzzifikálásának folytatása

- közepes nehézségű (KZ) *0.68* mértékben



- rosszul jegyzetelt (RJ) *0.48* mértékben



A példa fuzzy szabályai

- Vesszük a tanulási_időre következtető szabályokat:
  - HA tananyag_nehézsége=közepes (*KZN*)
    ÉS jegyzeteltség=rossz (*RJ*)
    AKKOR tanulási_idő=hosszú (*HI*)
  - HA tananyag_nehézsége=könnyű (*KNN*)
    ÉS jegyzeteltség=rossz (*RJ*)
    AKKOR tanulási_idő=közepes (*KI*)
- Mindkét szabály előfeltétele – bizonyos mértékben – illeszkedik a tényállításokhoz:
  - közepes nehézségű tananyag ($\mu_{KZN}(x)=0.68$)
  - rossz jegyzet ($\mu_{RJ}(x)$}$=0.48$)
  - könnyű tananyag ($\mu_{KNN}(x)= 0.22$)

Műveletek fuzzy állításokkal

- Alapműveletek: (A és B fuzzy halmazok)
  - $\mu_{\bar{A}}(x) = \neg\mu_A(x) = 1\text{-}\mu_A(x)$
  - $\mu_{A\cap B}(x) = (\mu_A \wedge \mu_B)(x) = \min\{\mu_A(x), \mu_B(x)\}$
  - $\mu_{A\cup B}(x) = (\mu_A \vee \mu_B)(x) = \max\{\mu_A(x), \mu_B(x)\}$
- Következtetés:

$$\frac{\mu_A(x) \text{ és } \mu_{A\subseteq B}(x,y)}{\mu_B(y) = \sup_{u\in U} \min\{\mu_A(x), \mu_{A\subseteq B}(x,y)\}} \qquad \frac{x\in A \text{ és ha } x\in A \text{ akkor } y\in B}{y\in B}$$

$$\frac{\mu_A(x) \text{ és } \mu_{A\subseteq B}(x)}{\mu_B(x) = \min\{\mu_A(x), \mu_{A\subseteq B}(x)\}} \qquad \frac{x\in A \text{ és ha } x\in A \text{ akkor } x\in B}{x\in B}$$

Fuzzy következtetés

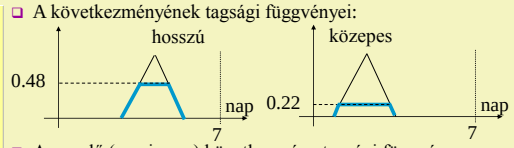
- Mindkét szabály előfeltételét kiszámoljuk
  - $(\mu_{KZN \cap RJ})(x) = \min\{\mu_{KZN}(x), \mu_{RJ}(x)\} = 0.48$
  - $(\mu_{KNN \cap RJ})(x) = \min\{\mu_{KNN}(x), \mu_{RJ}(x)\} = 0.22$
- Meghatározzuk a szabályok következményeit

$$\frac{\mu_A(x) \text{ és } \mu_{A \subseteq B}(x)}{\mu_B(x) = \min\{\mu_A(x), \mu_{A \subseteq B}(x)\}}$$

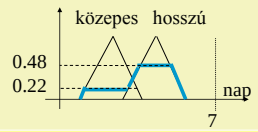
  - $\mu_{HI}(x) = (\mu_{KZN \cap RJ})(x) = 0.48$ mert $(\mu_{KZN \cap RJ \subseteq HI})(x)=1$
  - $\mu_{KI}(x) = (\mu_{KNN \cap RJ})(x) = 0.22$ mert $(\mu_{KNN \cap RJ \subseteq KI})(x)=1$

A példa következményeinek tagsági függvényei

- A következményének tagsági függvényei:

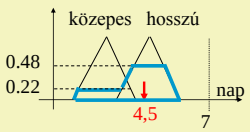


- Az eredő (maximum) következmény tagsági függvénye:



A példa defuzzifikálása

- A kapott eredmény számszerűsítése centroid (súlypont) módszerrel: vetítsük le az eredő tagsági függvény grafikonja és a vízszintes tengely által bezárt síkidom súlypontját a vízszintes tengelyre.



- Tehát *4,5* napig kell tanulni.

Megjegyzés

- Ellentétben a Bayes modellel, ahol a bizonytalansági mérték az események előfordulásának bizonytalanságát fejezi ki, itt a rosszul vagy pontatlanul definiált események előfordulásának mértékével foglalkozunk.
- Szemben heurisztikus modellek bizonyossági tényezőjével (MYCIN), amelyek azt fejezik ki, hogy mennyire vagyunk biztosak abban, hogy egy elem beletartozik-e egy halmazba, itt azzal foglalkozunk, hogy milyen mértékben tartozik bele.

Fuzzy modell értékelése

- Előnyei:
  - Közel áll az ember valóságszemléletéhez: nem kell számszerűsíteni a bizonyosság mértékét.
  - A numerikus modellekkel szemben egyszerűbb a rendszerleírás.
  - Előnyösen alkalmazható hiányos és bonyolult feladatok esetén
  - Könnyű vele számolni.
- Hátrányai:
  - Tagsági függvény megadása nehéz, nincs mögötte elméleti megalapozás.
  - Kombinációs függvények kritikája (pl. üres metszetű halmazok)

5. Dempster-Shafer elmélet

- Az állítások valószínűségének van egy bizonytalansága, amit a tények ismeretének hiánya okoz, hiszen egy állítás bekövetkezésének valószínűsége erősen függ attól, hogy milyen tényeket ismerünk.
- Mennyi annak a valószínűsége, hogy egy tény támogatja az állítást?
  - „Mit mondjunk, milyen valószínűséggel dobhatunk fejet egy pénzérmével, ha tudjuk, hogy az érme szabálytalan; mit, ha szabályos; vagy mit, ha 90%-os valószínűséggel szabályos?”

Valószínűség változása a tények függvényében

- Ha nincs bizonyíték nincs arra, hogy az érme szabályos:
  - Mennyire higgyük, hogy a dobás ½ valószínűséggel fej lesz? Semennyire: 0 * ½ = 0. Ez a fej-dobás valószínűségének alsó határa, a fej-dobásba vetett teljes bizalom.
  - Mennyire higgyük, hogy a dobás ½ valószínűséggel nem lesz fej? Semennyire. Tehát az ellenkező esemény, a fej-dobás valószínűségének felső határa, az elfogadhatósága: 1–0* ½ =1.
  - A fej-dobás kimenetelének bizonytalansága: [0,1]
- Ha az érme 90%-os valószínűséggel szabályos.
  - Fej-dobás valószínűségének alsó határa: 0.9* ½ = 0.45
  - Fej-dobás valószínűségének felső határa, 1–0.9* ½ = 0.55
  - A fej-dobás kimenetelének bizonytalansága: [0.45,0.55]
- Szabályos érme esetén az intervallum [½, ½]

Fogalmak

- Hiedelem függvény értékének számítása a valószínűségi alap függvény segítségével történik:
  - $Bel(a)=\Sigma_{x\subseteq\{a\}}\, m(x)$
- Bizalom intervallum: $[Bel(a), 1-Bel(\neg a)]$
- Valószínűségi alap függvény
  Legyen Θ elemi állítások halmaza. Adott tény(ek)hez tartozó $m:2^{\Theta}\rightarrow[0,1]$ függvény azt mutatja meg, hogy az $a\subseteq\Theta$ összetett állítás milyen mértékben ($m(a)$) függ a tényektől.
- Definíció szerint a valószínűségi alap függvényre teljesül
  - $\Sigma_{x\subseteq\Theta}\, m(x)=1$
  - $m(\emptyset)=0$

0. Példa

- Három fogoly (*A,B,C*) tudja, közülük egyet kivégeznek (H_X), a többieket felmentik.
  - $p(H_A)= p(H_B)= p(H_C)= ½$
- Ha semmilyen tényt nem ismerünk, akkor csak az biztos, hogy $\{H_A, H_B, H_C\}$ esemény biztosan bekövetkezik, azaz $m(\{H_A, H_B, H_C\})=1$. Azért, hogy $\Sigma_{x\subseteq\Theta}\, m(x)=1$ teljesüljön: $m(\{H_A\})=m(\{H_B\})=m(\{H_C\})=m(\{H_B,H_C\})= ...= 0$
- Ekkor *A*-nak a kivégzésébe vetett hite: **[0,1]**
  - $Bel(H_A) = \Sigma_{x\subseteq\{H_A\}}\, m(x) = m(\{H_A\})= 0$
  - $Bel(\neg H_A)= Bel(H_B,H_C) = \Sigma_{x\subseteq\{H_B\,H_C\}}\, m(x) =$
    $= m(\{H_B\})+ m(\{H_C\}) +m(\{H_B,H_C\}) = 0+0+0 =0$

1. Példa

- Az *A* fogoly megkéri az őrt, vigyen el egy levelet annak a fogolytársának, akit felmentettek. Az *A* megtudja, hogy levelet az őr *80*%-os valószínűséggel *B*-nek adta.
  - $p(H_B)= 0, p(H_A)= p(H_C)= ½$
  - $m(\{H_B\})= 0.8* 0, m(\{H_C\})= m(\{H_A\})=0.8*½$
  - $m(\{H_A, H_B, H_C\})= 0.2$ (hogy $\Sigma_{x\subseteq\Theta}\, m(x)=1$ teljesüljön)
- Ekkor *A*-nak a kivégzésébe vetett hite: **[0.4, 0.6]**
  - $Bel(H_A) = \Sigma_{x\subseteq\{H_A\}}\, m(x) = m(\{H_A\})= 0.4$
  - $Bel(\neg H_A)= Bel(H_B,H_C) = \Sigma_{x\subseteq\{H_B\,H_C\}}\, m(x) =$
    $= m(\{H_B\})+ m(\{H_C\})= 0 + 0.4= 0.4$

2. Példa

- Az *A* fogoly (most nem történt meg az előző példában említett levelezés) csak azt tudja *60*%-os bizonyossággal az őrtől, hogy *C* a halálraítélt.
  - $p(H_C)= 1, p(H_A)= p(H_B)= 0$
  - $m(\{H_C\})=0.6,\ m(\{H_A\}) = m(\{H_B\}) = 0,$
  - $m(\{H_A,H_B,H_C\})= 0.4$ (hogy $\Sigma_{x\subseteq\Theta}\, m(x)=1$ teljesüljön)
- Ekkor *A*-nak a kivégzésébe vetett hite: **[0, 0.4]**
  - $Bel(H_A) = \Sigma_{x\subseteq\{H_A\}}\, m(x) = 0$
  - $Bel(\neg H_A)= Bel(H_B,H_C)= \Sigma_{x\subseteq\{H_B\,H_C\}}\, m(x) =$
    $= m(\{H_B\})+ m(\{H_C\})= 0 + 0.6$

Eredő valószínűségialap függvény

- Eredő valószínűségi alap függvény:
$$m_E(z)= \frac{\Sigma_{x\cap y=z}\ m_1(x)m_2(y)}{1-\Sigma_{x\cap y=\emptyset}\ m_1(x)m_2(y)} \qquad z\neq\emptyset$$
- Matematikai alapok: Az eredő függvényre teljesülnek a valószínűségi alap függvény kritériumai.
  $\Sigma_{z\subseteq\Theta,\, z\neq\emptyset}\, m_E(z) = 1$
  hiszen
  $$\Sigma_{z\subseteq\Theta,\, z\neq\emptyset}\ \Sigma_{x\cap y=z}\ m_1(x)m_2(y) =$$
  $$= \Sigma_{z\subseteq\Theta}\Sigma_{x\cap y=z}\ m_1(x)m_2(y)-\Sigma_{x\cap y=\emptyset}\ m_1(x)m_2(y)=$$
  $$= \Sigma_{x\subseteq\Theta}\ \Sigma_{y\subseteq\Theta}\ m_1(x)m_2(y)-\Sigma_{x\cap y=\emptyset}\ m_1(x)m_2(y)=$$
  $$= \Sigma_{x\subseteq\Theta}\ m_1(x)\ \Sigma_{y\subseteq\Theta}\ m_2(y)-\Sigma_{x\cap y=\emptyset}\ m_1(x)m_2(y)$$
  $$= 1-\Sigma_{x\cap y=\emptyset}\ m_1(x)m_2(y)$$

Hiedelmek összevetése

$\Sigma_{x\cap y=z}\ m_1(x)m_2(y)$

m_2 \ m_1	$\{H_A\}$ 0.4	$\{H_C\}$ 0.4	Θ 0.2
$\{H_C\}$ 0.6	$\varnothing$ 0.24	$\{H_C\}$ 0.24	$\{H_C\}$ 0.12
Θ 0.4	$\{H_A\}$ 0.16	$\{H_C\}$ 0.16	Θ 0.08

$\Sigma_{x\cap y=\varnothing}\ m_1(x)m_2(y) = 0.24$

$m_E(\{H_A\}) = 0.16/(1-0.24) = 0.21$

$m_E(\{H_C\}) = (0.24+0.12+0.16)/(1-0.24) = 0.68$

$m_E(\{\Theta\}) = 0.08/(1-0.24) = 0.11$

1. és 2. példa összevonása

- Az *A* tudja, hogy az őr *80*%-os valószínűséggel *B*-nek vitte a levelet, és *60*%-os valószínűséggel az őr szerint *C*-t ítélték halálra.
- A két tény valószínűségi alap függvényéből kiszámoltuk a két tényre együttesen támaszkodó eredő valószínűségi függvényt. Ez alapján a hiedelem intervallum: **[*0.21,0.32*]**
  - $m(\{H_A\})=0.21, m(\{H_C\})=0.68, m(\{H_A,H_B,H_C\})=0.11$
  - $Bel(H_A) = \Sigma_{x\subseteq\{H_A\}}\ m(x) = m(\{H_A\}) = 0.21$
  - $Bel(\neg H_A) = Bel(H_B,H_C) = \Sigma_{x\subseteq\{H_B\,H_C\}}\ m(x) = m(\{H_B\}) + m(\{H_C\}) = 0 + 0.68$

Dempster-Shafer elmélet értékelése

- Előnyei:
  - A klasszikus valószínűség számítással szemben nem elemi eseményekkel foglalkozik
  - Nem igényli apriori valószínűségek megadását
  - Matematikailag jól megalapozott
  - Képes különböző tények bizonyító erejét összegezni.
- Hátrányai:
  - Összes megfigyelhető esetet elő kell állítani,
  - amelyekre az alap valószínűségi függvény értékeket meg kell adni.

MYCIN értékelése

- Egyszerűsített változataival több korai szakértő rendszernél találkozhatunk.
  - M.1 bizonyossági faktora 0 100 közötti (százalékos) értéket ad és közvetlenül kombinálható.
- Elméleti háttér nélkül (a mértékek nem mutatnak valószínűségi eloszlást) bizonyos feladatokon meglepően jól használható.
- Kevés a priori érték tárolása mellett kicsi számításigényű kalkulus.
- Jól illeszkedik a szabályalapú következtetéshez, a következmény bizonyossági mértéke leválasztható a feltételekről.