AMI ZH **60 perc**
2020. nov. 20.

PótZH

∑

Max. 30 pont **Név (nyomtatott betűkkel):** ______________________________

Szükséges minimum: 12 pont

Neptun-kód:

Meg nem engedett segédeszközt vagy segítséget nem vettem igénybe.

____
aláírás

Feladat sorszáma	**1**	**2**	**3**	**4**	**5**	**6**
Kapott pontok						

---

1. Az alábbi állításoknál a helyes választ (IGAZ/HAMIS) kell bekarikázni. Minden jó válasz +1 pont, minden rossz válasz -0,5 pont (a nem megválaszolt kérdés értelemszerűen 0 pont). Ha negatív lenne a végső pontszám ebben a feladatban, akkor nullára „kerekítjük”.

12 p. ______

a. A mélységi keresés időigénye lehet nagyobb a szélességi keresésnél. a. IGAZ HAMIS

b. Az első attribútum értékének tesztjére alapozott döntési eljárást jellemző ROC görbe alatti terület (AUC) 0,75; a második attribútum értékének tesztjére alapozott döntési eljárásra AUC 0,83. Ennek alapján a második attribútum tesztjét tekintjük jobbnak. b. IGAZ HAMIS

c. Ha az *n*-edik csomópontra a kezdeti állapottól eddig összegyűlt költségünk, *g*(*n*) kisebb, mint az *m*-edik csomópontra adott *g*(*m*) érték, akkor az egyenletes költségű keresési algoritmus biztosan a *n*-ediket fejti ki először. c. IGAZ HAMIS

d. A mohó keresés mindig optimális megoldást talál. d. IGAZ HAMIS

e. Egy kétosztályos döntési probléma bemeneti leírására 10 bináris attribútumot használunk. A döntési fa egyik bejárási útján már mind a 10 attribútumot teszteltük. Ez esetben a kialakuló levélcsomópont biztosan homogén lesz (vagy csupa pozitív vagy csupa negatív példa jut ide). e. IGAZ HAMIS

f. Azzal kerestek meg, hogy szabályalapú rendszerünkkel bizonyítsunk be egy megadott tételt (sejtést). Ez esetben tipikusan hátrafele következtetést használunk. f. IGAZ HAMIS

g. Egy kétosztályos döntésre kialakított döntési fa egyik csomópontjában az információszükséglet nulla, ha mindkét osztályból ugyanannyi minta jut ebbe a csomópontba. g. IGAZ HAMIS

h. Egy teszt érzékenysége (TPR) mindig kisebb a specificitásánál (TNR). h. IGAZ HAMIS

i. Előrefele következtetésnél a szabályok feltételrészére igyekszünk illeszteni. i. IGAZ HAMIS

j. A valószínűségi hálók a változók közti feltételes függetlenségek kihasználásával adnak egyszerűbb, jobban kezelhető leírást a problémára. j. IGAZ HAMIS

k. Egy keresési fában a fa több csomópontjához is tartozhat a megoldani kívánt problémának ugyanaz az állapota k. IGAZ HAMIS

l. Az együttes eloszlás birtokában minden kérdésre választ kaphatunk, ami a benne szereplő véletlen változók viszonyára vonatkozik. l. IGAZ HAMIS

2. Egy B betegségben a lakosság 10%-a szenved. A B betegségben szenvedők 70%-a köhög. Ugyanakkor más betegségek is okoznak köhögést, így a lakosság 23%-a köhög.
Ha egy paciens köhögéssel jelentkezik az orvosnál, akkor mi a valószínűsége, hogy a B betegségben szenved? (Természetesen számítással és rövid 1-2 mondatos magyarázattal indokolja válaszát!)

Megoldás:

Az egész feladat tulajdonképpen a Bayes-tétel felírását jelenti. Jelöljük K-val a köhögést, B-vel a betegséget. 2 p. ______

$$P(K) = 0,23$$
$$P(B) = 0,1$$
$$P(K|B) = 0,7$$

A Bayes-tétel értelmében:

$$\mathrm{P}(\mathrm{B}|K) = \frac{P(K|B)\cdot P(B)}{P(K)} = \frac{0,7\cdot 0,1}{0,23} = 0,304 \quad (30,4\%)$$

3. Az ítéletlogika 7 általános következtetési szabálya közül nevezzen meg és írjon fel 3-at! (Nem kell magyarázat.)

3 p. ______

Általános következtetési szabályok

1. Modus Ponens $\dfrac{A \quad A\to B}{B}$

2. ÉS kiküszöbölés $\dfrac{A1 \wedge A2 \wedge A3 \wedge \ldots \wedge AN}{A1 \quad A2 \quad \ldots \quad AN}$

3. ÉS bevezetés $\dfrac{A1 \quad A2 \quad \ldots \quad AN}{A1 \wedge A2 \wedge A3 \wedge \ldots \wedge AN}$

4. Dupla negálás $\dfrac{\neg(\neg A)}{A}$

5. VAGY bevezetés $\dfrac{Ak}{A1 \vee A2 \vee \ldots \vee \color{red}{Ak} \vee \ldots \vee AN}$

6. Egységrezolúció $\dfrac{\neg A \quad A\vee B}{B}$

7. Rezolúció $\dfrac{A\vee B \quad \neg B\vee C}{A\vee C}$

4. Egy tesztet végzünk, leteszteljük a teljes 4 milliós lakosságot. A teszt érzékenysége (TPR: True Positive Ratio) 85%, a specificitása (TNR: True Negative Ratio) 96%. Ha tudjuk, hogy a 4 milliós lakosság 5%-a fertőzött, akkor hány fertőzöttnél és hány egészségesnél fog a teszt betegséget (fertőzöttséget) jelezni? A teszt által fertőzöttnek *jelzett* emberek hány százaléka valójában egészséges? (Rövid magyarázat és számítás szükséges!)

Megoldás:

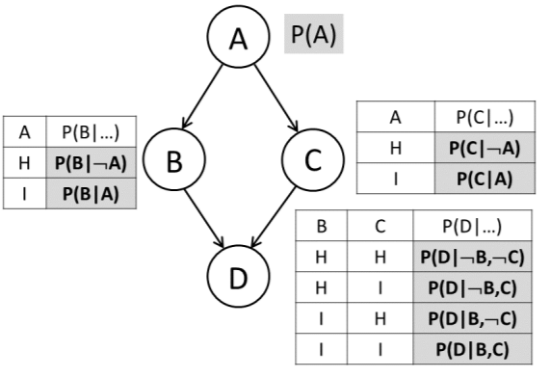
A 4.000.000 lakosból 5% fertőzött, ez 4.000.000*0,05=200.000 ember. 3 p. ______

A 200.000 fertőzöttből 0,85*200.000=170.000-nél mutatja ki a teszt a fertőzöttséget.

A 4.000.000-200.000=3.800.000 egészségesből 96%-ot helyesen felismer egészségesnek, de a maradék 4%-ot fertőzöttnek jelzi a teszt (fals pozitív). Ez 3.800.000*0,04=152.000 ember.

Tehát a teszt összesen 170.000+152.000=322.000 fertőzöttet fog jelezni, de ezek közül 152.000/322.000=0,472 azaz (47,2%) valójában egészséges.

5. Adott a következő valószínűségi háló. 4 p. ______



Írja fel annak valószínűségét, hogy B IGAZ értékű, feltéve, hogy A és D IGAZ, de C HAMIS értéket vesz fel! A keresett feltételes valószínűséget ki kell fejezni úgy, hogy a végeredményben *kizárólag az ábrán megadott valószínűségek szerepelnek*! (Pl. a P(¬X)-et ki kell fejezni P(X)-el, ha csak az utóbbi van megadva stb.) Számértékeket nem adtunk meg, mert nem kell kiszámítani. (Természetesen 1-2 mondatos magyarázat, rövid levezetés vagy magyarázó ábra kell.)

Megoldás:

A keresett feltételes valószínűség: P(B|A,¬C,D). A Bayes-tételből indulunk ki. Az alábbi megoldásnál persze lehet egyszerűbbet alkotni (A P(¬C|A)-ra nincs szükség, hiszen C értéke rögzített stb.)

$$P(B|A,\neg C,D)=\frac{P(B,A,\neg C,D)}{P(A,\neg C,D)}=\frac{P(B,A,\neg C,D)}{\sum_b P(b,A,\neg C,D)}=$$

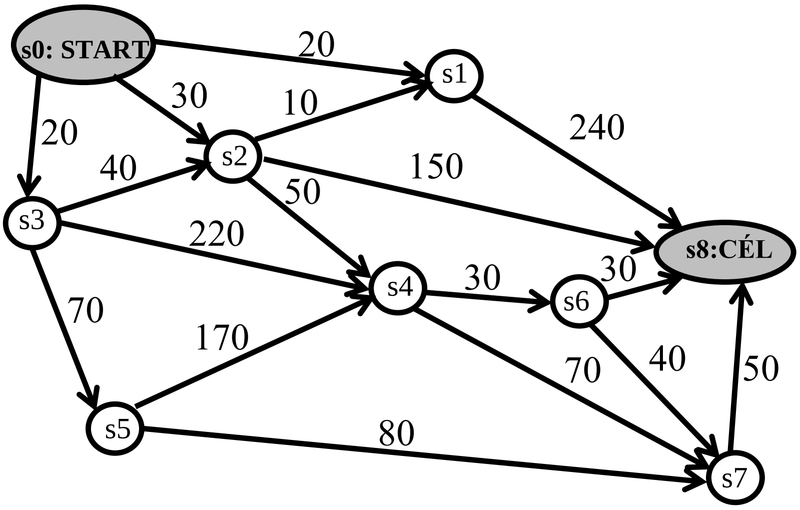
$$=\frac{P(B,A,\neg C,D)}{P(B,A,\neg C,D)+P(\neg B,A,\neg C,D)}=$$

$$=\frac{P(A)\cdot P(B|A)\cdot P(\neg C|A)\cdot P(D|B,\neg C)}{P(A)\cdot P(B|A)\cdot P(\neg C|A)\cdot P(D|B,\neg C)+P(A)\cdot P(\neg B|A)\cdot P(\neg C|A)\cdot P(D|\neg B,\neg C)}=$$

$$\frac{P(A)\cdot P(B|A)\cdot (1-P(C|A))\cdot P(D|B,\neg C)}{P(A)\cdot P(B|A)\cdot (1-P(C|A))\cdot P(D|B,\neg C)+P(A)\cdot (1-P(B|A))\cdot (1-P(C|A))\cdot P(D|\neg B,\neg C)}=$$

$$=\frac{P(B|A)\cdot P(D|B,\neg C)}{P(B|A)\cdot P(D|B,\neg C)+(1-P(B|A))\cdot P(D|\neg B,\neg C)}$$

6. Az alábbi állapotokkal és lehetséges egyirányú állapotátmenetekkel jellemzett problémát A* kereséssel oldjuk meg. (Mivel egyirányúak az átmenetek, soha nem lépünk vissza abba az állapotba, ahonnan érkeztünk.) Az ábrán feltüntettük az állapotátmenetek költségét, a mellékelt táblázat mutatja a heurisztikánk egyes állapotokhoz tartozó értékét: 6 p. ____

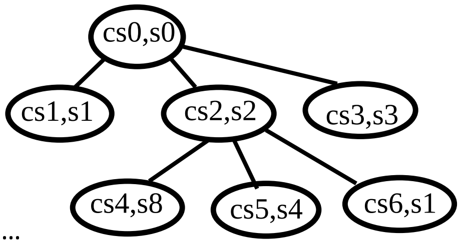


állapot (*n*)	*h*(*n*)
s0	**140**
s1	**200**
s2	**110**
s3	**130**
s4	**50**
s5	**120**
s6	**15**
s7	**40**
s8	**0**

A keresés két listát épít, az elsőben (closedL) azok a csomópontok szerepelnek, amiket már kifejtett, a másodikban azok, amelyekhez már eljutott, de még nem fejtette ki ezeket (openL). Mindegyik listaelem 5 mezőből épül fel (rövidítés: csp.=csomópont):

(szülőcsp., aktuális csp., állapot, eddig megtett út költsége, az aktuális csp.-hoz a heurisztika értéke),

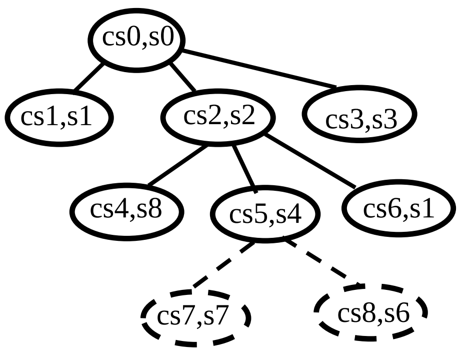
például a gyökércsomópontra: (-,cs0,s0,0,140).



A két lista a második lépés után:
closedL={(-,cs0,s0,0,140), (cs0,cs2,s2,30,110)}
openL={(cs2,cs5,s4,80,50), (cs0,cs3,s3,20,130), (cs2,cs4,s8,180,0) ,…
...., (cs0,cs1,s1,20,200), (cs2,cs6,s1,40,200) }

Adja meg a következő lépés után kialakuló keresési gráfot és a két listát! (Itt nem kell külön indoklás!)

Megoldás: (szaggatottal az új rész)



A két lista a következő lépés után:

closedL={(-,cs0,s0,0,140), (cs0,cs2,s2,30,110), **(cs2,cs5,s4,80,50)**}
openL={**(cs5,cs8,s6,110,15)**, (cs0,cs3,s3,20,130), (cs2,cs4,s8,180,0) , **(cs5,cs7,s7,150,40)**, ...
…………………………………………………………………………………. (cs0,cs1,s1,20,200), (cs2,cs6,s1,40,200) }