

A1. Mi az adatbányászat alapvető célkitűzése? (vendégelőadás) (2 pont)

A2. Fennáll-e az alábbi vonzat: $(A \wedge B) \models (A \Leftrightarrow B)$?

A választ az igazságtábla alapján modellvizsgálattal fejtse ki! Miért igen/ miért nem? Mi jelent az, hogy vonzata? (2 pont)

A3. Adva van két tetszőleges, de elfogadható $h1$ és $h2$ heurisztika. Melyik összetett $\max(h1, h2)$, $(h1 + h2)/2$, ill. $\min(h1, h2)$? heurisztikát ajánlatos alkalmazni és miért? (3 pont)

A4. Mire szolgál a korlátozáskielégítési problémáknál az élkonzisztencia lépés? Mi a lényege (egy példával megtoldva)? (3 pont)

A5. A közgazdászok a rizikót kerülő ember pénzhasznosság érzetét sokszor exponenciális $U(x) = R * (1 - \exp(-x/R))$ képlettel modellezik, ahol R egy pozitív konstans (x -szel azonos dimenziójú), amely az egyén rizikótűrését fejezi ki. Tegyük fel, hogy Béla hasznosság képletében $R = 500$ Ft. Béla egy fix 1000 Ft összeget fogad egy olyan sorsjáték ellentételezéseként, ahol p valószínűséggel 1 millió Ft-t kaphat, de $(1-p)$ valószínűséggel 0-t. Mi lehet a p valószínűség értéke? (5 pont)

Vegye figyelembe, hogy: $\exp(-10) = (\text{kb.}) 0$, $\exp(-2) = .1353$

A6. Mi a tanulásnál felmerülő '(tanulási) torzítás' fogalma? Miért lényeges? (2 pont)

A7. Tegyük fel, hogy egy drog teszt 98% pontos a droggal élő személy tesztelése esetén (azaz a teszt .98 valószínűséggel pozitív, ha a tesztelt személy droggal él) és 90% pontos a drogot nem használó személy tesztelése esetén (azaz a teszt .9 valószínűséggel negatív, ha a tesztelt személy nem él droggal).

Azt is tegyük fel, hogy az egész populáció 10%-a használ drogot. Egy személy jelentkezik a tesztre és annak eredménye pozitív. Mi a valószínűsége, hogy a tesztelt személy ténylegesen él droggal?

Mi ennek a tesznek a hamis pozitív valószínűsége (hogy a droggal nem élő valaki bizonyul teszt-pozitívnak)? Mi ennek a tesznek a hamis negatív valószínűsége (hogy a droggal élő valaki bizonyul teszt-negatívnak)?

Legyenek az alábbi jelöléssel:

P = az az esemény, hogy a teszt pozitívnak bizonyult valaki számára,

N = $(\neg P)$ az az esemény, hogy a teszt negatívnak bizonyult valaki számára,

A = az az esemény, hogy a tesztelt személy tényleg drogfogyasztó.

(4 pont)

A8. Arisztotelész egyik szillogizmusa modern átírásban:

CAMESTRES:

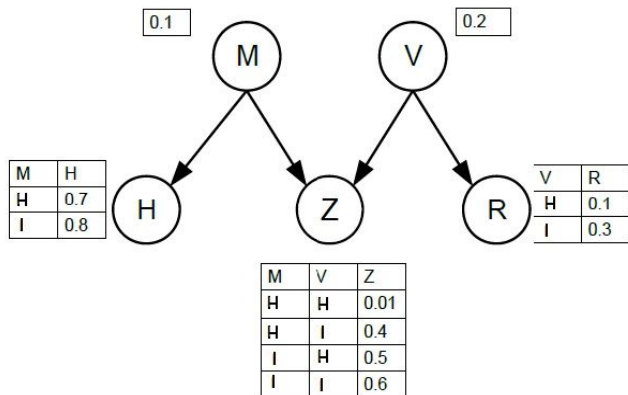
$$\forall x. B(x) \rightarrow A(x)$$

$$\forall x. C(x) \rightarrow \neg A(x)$$

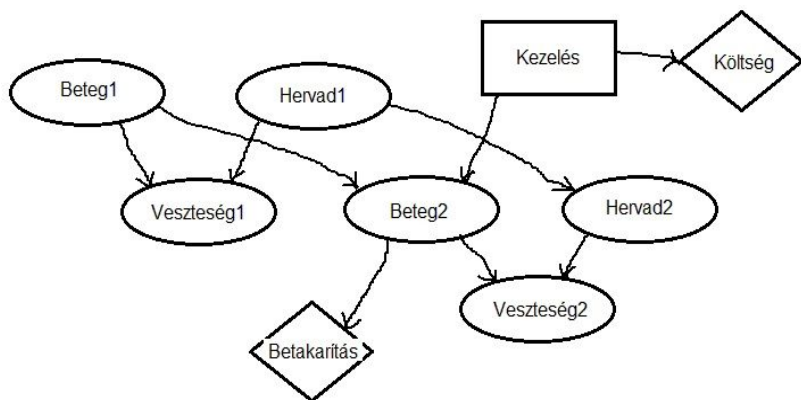
$$\forall x. C(x) \rightarrow \neg B(x)$$

Lássuk be a rezolúciós bizonyítással, vajon igaza volt-e! (8 pont)

A9. Számítsa ki a háló alapján, hogy mi a $P(M | R)$ valószínűség értéke? És mi a $P(M | R, Z)$ valószínűség értéke? (6 pont)



A10. Az alábbi döntési háló almaszüretet modellez. Alma betegsége, ill. elhervadása csökkenti a jó betakarítás lehetőségét. A háló 1 hónappal eltolt helyzetet ábrázol, amikor még észbe kapva kezelhetjük az almafákat, a betakarítás kézbetartása érdekében. Mindegyik véletlen változó bináris (0/1, Kicsi/Nagy), ill. a Kezelés lehet: Nem vagy Igen. Vizsgáljuk meg, vajon érdemes-e a fákat kezelni, ha Hervad1 = Nagy? (5 pont)



Legyen $P(\text{Beteg1}) = P(\text{Hervad1}) = .1$. Tegyük fel továbbá, hogy a csomóponti valószínűségek:

Beteg1	Hervad1	P(Veszteség1)
0	0	.02
0	1	.85
1	0	.9
1	1	.95

Beteg1	Kezel	P(Beteg2)
0	N	.02
0	I	.01
1	N	.99
1	I	.2

Hervad1	P(Hervad2)
0	.05
1	.60

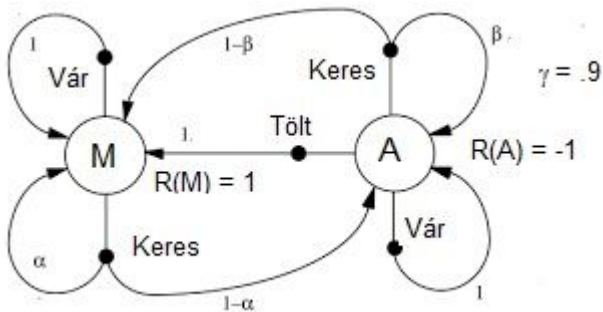
A $P(\text{Veszteség2} \mid \text{Beteg2}, \text{Hervad2})$ tábla azonos az '-1'-es végződésű változókéval.

A hasznosságok kifejezése:

$$U(\text{Költség}) = -8000 \times (\text{Kezel} = \text{Igen}) + 0 \times (\text{Kezel} = \text{Nem})$$

$$U(\text{Betakarítás}) = 3000 \times P(\text{Beteg2}) + 20000 \times P(\neg\text{Beteg2})$$

A11. Az alábbi szekvenciális döntési feladatnál egy robotnak két (M és A akku feltöltési) állapota van, ill. 3 cselekvéssel rendelkezik (Vár, Keres, Tölt). Adja meg és oldja meg a robot optimális eljárás módjához tartozó hasznosságok Bellmann egyenletét! (5 pont)



Tipp: A Bellmann egyenlet megoldásánál vizsgálja meg, melyik $U(M) < U(A)$, vagy $U(M) > U(A)$ esetről lehet egyáltalán szó?

A12. Starbucks kutatói vendégekről egy adatbázist gyűjtettek be (alábbi 10 példa). Kávénak két ízesítése lehet (Mogyoró hozzáadva = Igen, stb.), ill. három állapotú pörköléserese (L alacsony, M közepes, D sötét). Ha a kutatók a vendégek ízlését egy döntési fával szeretnék kifejezni, melyik attribútummal kellene kezdeni a fa építését? (5 pont)

Sz.	Mogyoró	Vanília	Pörkölés	Kedvel
X1	I	I	L	I
X2	I	I	M	I
X3	I	I	D	I
X4	I	I	L	I
X5	I	N	M	N
X6	I	N	D	N
X7	N	N	L	N
X8	N	I	M	N
X9	N	N	D	I
X10	N	N	L	I

Vegye figyelembe, hogy: $I(2/5, 3/5) = .9710$, $I(1/3, 2/3) = .9183$, $I(3/8, 5/8) = .9544$, $I(1/5, 4/5) = .7219$, $I(1/4, 3/4) = .8113$.