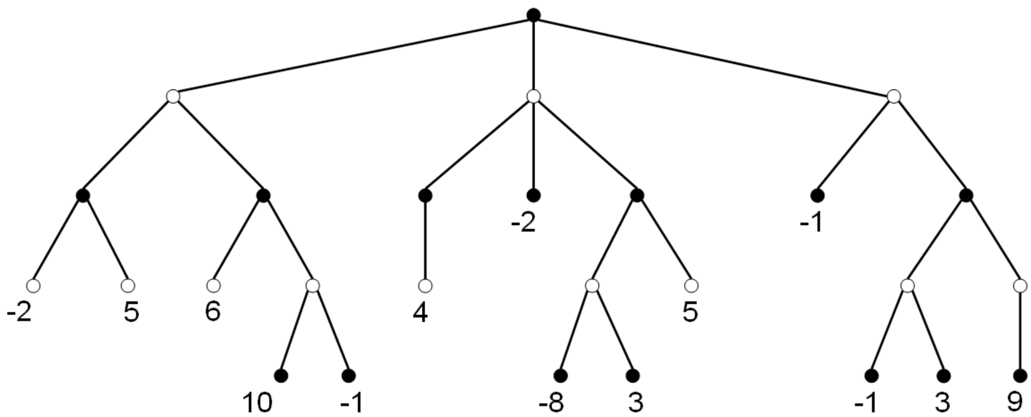
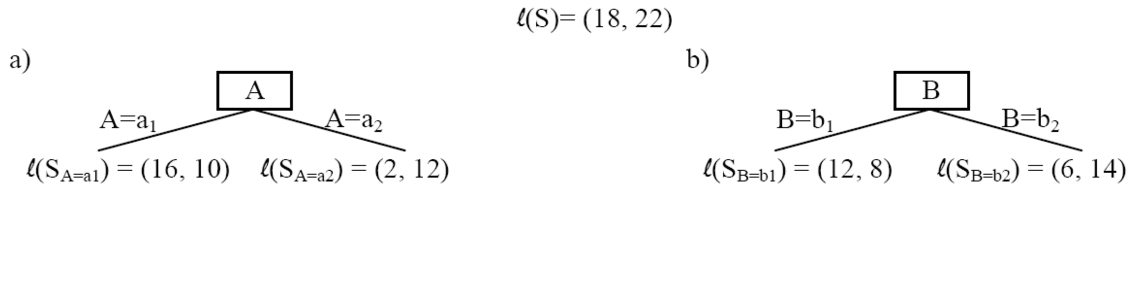
1. A lenti táblázat egy kereskedelemmel foglalkozó vállalat adatbázisa, amit a hitelfelvétel előtt álló személyekről és azok hitelképességéről – miután a bank elbírálta azt – gyűjtött össze. Az adatbázis segítségével *hozz Naive-Bayes döntést* arról, hogy egy személy, aki **idős** korú, **magas** bevétellel rendelkező, **nem** tanuló valamint **nincs** megtakarítása valószínűleg hitelképesnek lesz-e nyilvánítva? Jelöld, hogy milyen valószínűségeket számolsz, mennyi azok értéke és azokból hogy hozol döntést.

ID	Kor	Bevétel	Tanuló-e	Befektetés	Hitelképes?
1	fiatal	magas	nem	nincs	Nem
2	fiatal	magas	nem	van	Nem
3	közép	magas	nem	nincs	Igen
4	idős	közepes	nem	nincs	Igen
5	idős	alacsony	igen	nincs	Igen
6	idős	alacsony	igen	van	Nem
7	közép	alacsony	igen	van	Igen
8	fiatal	közepes	nem	nincs	Nem
9	fiatal	alacsony	igen	nincs	Igen
10	idős	közepes	igen	nincs	Igen
11	fiatal	közepes	igen	van	Igen
12	közép	közepes	nem	van	Igen
13	közép	magas	igen	nincs	Igen
14	idős	közepes	nem	van	Nem

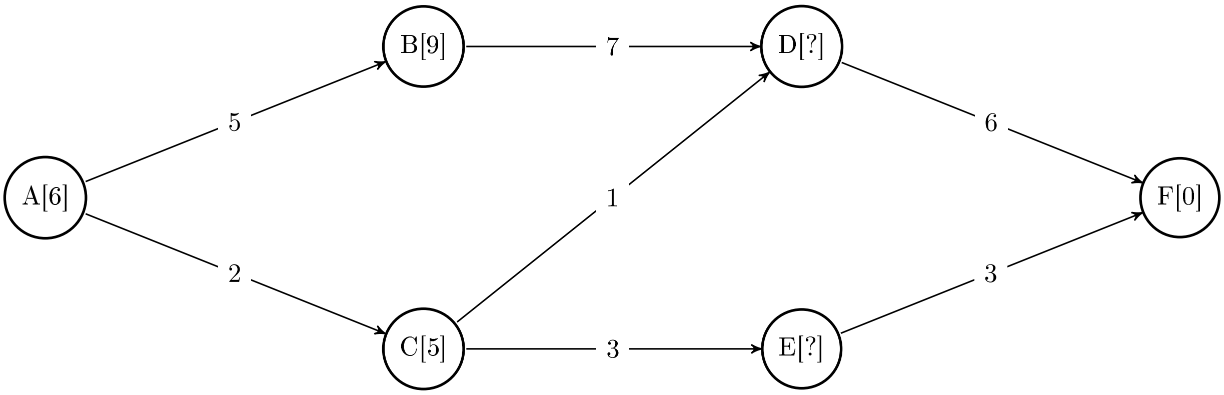
2. Hajtsd végre a *MinMax-$\alpha\beta$ algoritmust ($\alpha-\beta$ vágást)* a lenti játékfán. A gyökér csomópont a *max* játékoshoz tartozik.



3. Az ID3 futása közben a konstruálás alatt álló fa egyik csúcsánál a lenti ábrán látható döntési helyzet adódik – azaz a csúcs vagy A címkét kap, és ezáltal az a.) eset áll elő, vagy B címkét kap, és ezáltal a b.) helyzet áll elő. A Shannon-féle entrópia helyett a Gini-féle $E(S) = 4p_+p_-$ indexet használva melyik attribútumot választja az ID3, és miért? (S a csúcshoz tartozó példák halmaza; $S_{F=f}$ azon S-beli példák halmaza, amelyeknél az F attribútum f értéket vesz fel; p_+ a pozitív példák, p_- a negatív példák előfordulási valószínűsége S-ben; ℓ első komponense pedig mindig a paraméterében lévő példahalmaz pozitív, a második pedig a negatív osztálycímkével rendelkező elemeinek száma.)

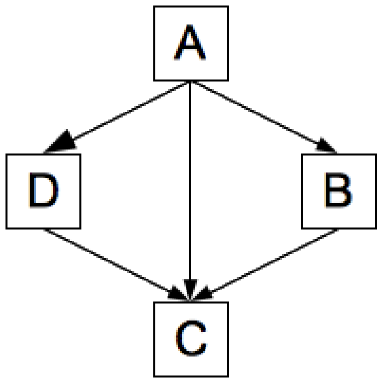


4. Az alábbi ábrával illusztrált állapottérben hajtsd végre az A* algoritmust (A - kezdőállípot, F - végállapot), valamint rendelj a D és E csúcsokhoz olyan – az algoritmus helyességéhez szükséges, tehát a monotonitást/konzisztenciát nem sértő – heurisztikus függvényértéket, hogy a két csúcs közül a D jelű hamarabb kerüljön a zárt halmazba az algoritmus helyes végrehajtása során!



5. Az ábrán látható Bayes háló alapján alakítsd olyan alakra az alábbi valószínűséget, amely értékei kiolvashatók a valószínűségi táblákból:

P(B=i,C=h)=?



$$P(A = i) = 0,25$$

x	$P(B = i \mid A = x)$
i	0,5
h	0,8

x	$P(D = i \mid A = x)$
i	0,22
h	0,76

x	y	z	$P(C = i \mid A = x, B = y, D = z)$
i	i	i	0,5
i	h	i	0,1
h	i	i	0,2
h	h	i	0,0
i	i	h	0,15
i	h	h	0,33
h	i	h	0,2
h	h	h	0,05

Mesterséges Intelligencia I. Minta ZH megoldások

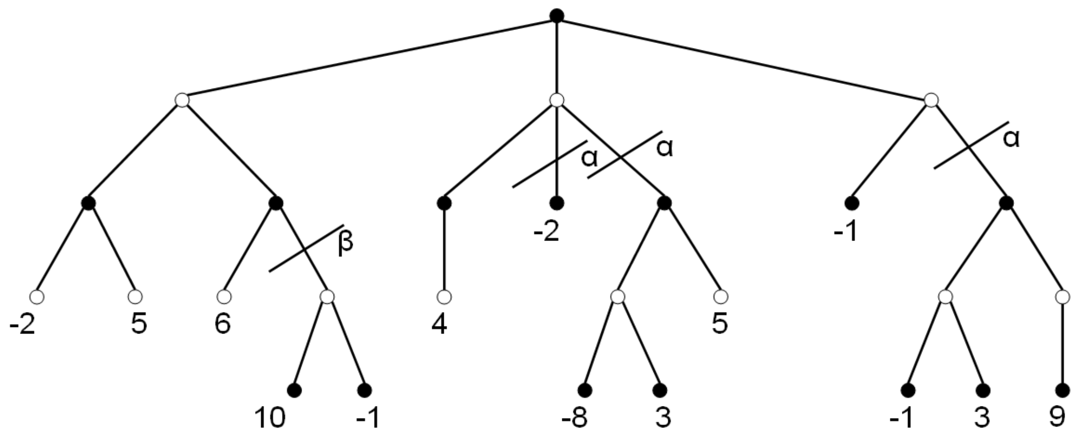
1. Naive-Bayes

   Az alábbi feltételes valószínűségek szorzata alapján hozok döntést.

$$P(N)P(i|N)P(m|N)P(n|N)P(n|N) = \frac{5}{14} \cdot \frac{2}{5} \cdot \frac{2}{5} \cdot \frac{4}{5} \cdot \frac{2}{5} = \frac{16}{7 \cdot 125}$$
$$P(I)P(i|I)P(m|I)P(n|I)P(n|I) = \frac{9}{14} \cdot \frac{3}{9} \cdot \frac{2}{9} \cdot \frac{3}{9} \cdot \frac{6}{9} = \frac{2}{7 \cdot 27}$$
$$\frac{16}{125} \cdot \frac{27}{2} = \frac{8 \cdot 27}{125} > 1$$

   A „Nem” osztály valószínűsge a nagyobb, így az előzetes vizsgálat alapján nem lesz hitelképes a vizsgált személy.

2. $\alpha - \beta$



   A fenti ábrán látható helyeken vannak vágások, a max játékos legalább 5 pontot szerez (ez az érték kerül a gyökér csomópontba).

3. ID3

$$Entropy(S) = 4 * \frac{18}{40} * \frac{22}{40} = \frac{99}{100}$$

$$Gain(S,A) = \frac{99}{100} - \left(\frac{26}{40} * 4 * \frac{16}{26} * \frac{10}{26} + \frac{14}{40} * 4 * \frac{2}{14} * \frac{12}{14}\right) = \frac{99}{100} - \left(\frac{8}{13} + \frac{6}{35}\right) = \frac{99}{100} - \frac{358}{455} \approx 0{,}20$$

$$Gain(S,B) = \frac{99}{100} - \left(\frac{20}{40} * 4 * \frac{12}{20} * \frac{8}{20} + \frac{20}{40} * 4 * \frac{6}{20} * \frac{14}{20}\right) = \frac{99}{100} - \left(\frac{12}{25} + \frac{21}{50}\right) = \frac{99}{100} - \frac{9}{10} = 0{,}09$$

   Mivel az A jellemzőhöz tartozó Gain-érték a nagyobb, emiatt az A jellemző mentén vágunk, tehát az a.) eset áll elő.

4. A*

Kezdetnek határozzuk meg $h(D)$ és $h(E)$ értékeket!

A heurisztika megengedőségének érdekében a heurisztikának nem szabad semelyik csúcsra sem felülbecsülni a célállapot eléréséig ténylegesen hátra lévő minimálisan szükséges távolságot, ezért $h(D) \leq 6$, illetve $h(E) \leq 3$ egyenlőtlenségeknek teljesüliük kell.

A heurisztika konzisztenciájának fennmaradása érdekében – azaz, hogy minden (n, n') irányított él mentén továbbra is fennálljon a $h(n) \leq cost(n, n') + h(n')$ egyenlőtlenség – pedig $h(D) \geq 4$ és $h(E) \geq 2$ egyenlőtlenségeknek kell teljesülniük, azaz tudjuk, hogy $4 \leq h(D) \leq 6$, valamint $2 \leq h(E) \leq 3$.

Végül ahhoz, hogy a feladatban szereplő további feltétel (D előbb kerüljön a zárt halmazba E-nél) teljesüljön, $f(D) = g(D) + h(D) = 3 + h(D) \leq 5 + h(E) = g(E) + h(E) = f(E)$-nek is teljesülnie kell. A heurisztikus függvény D és E pontokban vett értékeire tett korábbi megszorításaink alapján ez az egyenlőtlenség $h(D) = 4$ és $h(E) = 3$ választása mellett teljesül, ekkor ugyanis $f(D) = 3 + 4 \leq 5 + 3 = f(E)$.

Ny	Z
A(NULL, 0, 6)	
B(A, 5, 5+9=14); **C(A, 2, 2+5=7)**	A(NULL, 0, 6)
B(A, 5, 14); **D(C, 3,** $3 + h(D)$**)**; E(C, 5, $5 + h(E)$)	A(NULL, 0, 6); C(A, 2, 7)
B(A, 5, 14); **E(C, 5,** $5 + h(E)$**)**; F(D, 9, 9)	A(NULL, 0, 6); C(A, 2, 7); D(C, 3, $3 + h(D)$)
B(A, 5, 14); **F(E, 8, 8)**	A(NULL, 0, 6); C(A, 2, 7); D(C, 3, $3 + h(D)$); E(C, 5, $5 + h(E)$)
B(A, 5, 14)	A(NULL, 0, 6); C(A, 2, 7); D(C, 3, $3 + h(D)$); E(C, 5, $5 + h(E)$); F(E, 8, 8)

5. Bayes háló

$$
\begin{aligned}
P(B = i, C = h) &= \sum_{X \in \{i,h\}} P(C = h, B = i, D = X) \\
&= \sum_{X \in \{i,h\}} \sum_{Y \in \{i,h\}} P(C = h, B = i, D = X, A = Y) \\
&= \sum_{X \in \{i,h\}} \sum_{Y \in \{i,h\}} P(C = h | B = i, D = X, A = Y) \\
&\quad P(B = i, D = X, A = Y) \\
&= \sum_{X \in \{i,h\}} \sum_{Y \in \{i,h\}} P(C = h | B = i, D = X, A = Y) \\
&\quad P(B = i | A = Y) P(D = X, A = Y) \\
&= \sum_{X \in \{i,h\}} \sum_{Y \in \{i,h\}} P(C = h | B = i, D = X, A = Y) \\
&\quad P(B = i | A = Y) P(D = X | A = Y) P(A = Y)
\end{aligned}
$$

A fenti alak minden egyes valószínűsége kiolvasható a Bayes háló valószínűségi táblázataiból.