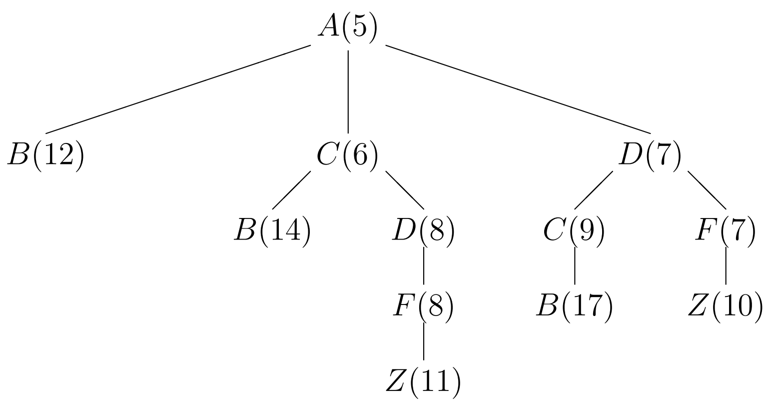
Uitwerkingen opgaven Kunstmatige intelligentie

Opgave 8 (1.6.2001)
e. Uiteindelijk wordt onderstaande boom opgebouwd. Knopen die al eerder op een pad voorkwamen worden niet aangegeven, er zijn dus geen kringen (bijvoorbeeld onder $C(6)$ komt niet opnieuw A te hangen). Tussen haakjes staat naast de knopen hun f-waarde. Achtereenvolgens worden geopend = ontwikkeld = geexpandeerd $A(5)$, $C(6)$, $D(7)$, $F(7)$, $D(8)$ (let op, hoewel al een doelknoop gevonden is, worden eerst nog knopen met lagere f-waarde ontwikkeld), $F(8)$, $C(9)$ en tot slot $Z(10)$. Bij $F(8)$ staat dankzij de *pathmax equation* 8 in plaats van 7, bij $F(7)$ staat evenzo 7 in plaats van 6.



f. De effectieve vertakkingsgraag b^* laat zich berekenen uit $1 + b^* + (b^*)^2 + \cdots + (b^*)^d = N$, met $d = 3$ de diepte van de oplossing, en $N = 8$ het aantal geopende knopen.

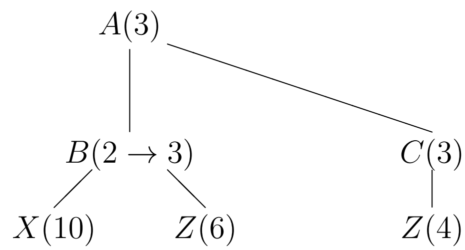
Opgave 9 (25.6.2003)
c. De route langs de D's kost $3k$, die langs de C's $2k + 3$ en die langs de B's kost $k + 8$. Dus voor $k < 3$: D-route, $k = 3$: C of D, $k = 4$: C, $k = 5$: C of B, en $k > 5$: B-route.
d. Veel verschillende f-waarden, zodat je vaak opnieuw een depth-first wandeling moet maken — slecht. Veelvuldig pathmax toepassen is fijn.

Opgave 10 (4.6.2006)
c. Vanuit F gezien: $1 \leq \sigma$; vanuit C gezien: $3 \leq 2\sigma$; en vanuit A gezien: $5 \leq 2 + 2\sigma$. In de andere knopen wordt automatisch onderschat. Neem dus $\sigma \geq 3/2$ (dan admissibel).
d. Bij het ontwikkelen van C–5 kom je F–$3 + \sigma$ tegen. Als $\sigma \leq 2$ wordt dit dankzij "pathmax" natuurlijk 5. Let er tevens op dat als $2 < \sigma < 3$ bij f-limiet 6 ook knoop F ontwikkeld wordt (anders alleen B).

Opgave 11 (21.8.2006)
b. Drie wijzigingen: $H(5)$, $D(5)$ en $E(6)$.
c. Bij F pathmax: $3 \to 5$. Als $\rho < 1$ krijg je $A - H - Y$ met $5 + \rho$ als afstand. Als $\rho > 1$ krijg je $A - D - G - Z$ met gewicht 6. Als $\rho = 1$ (een van) beide.
d. De beste admissibele waarde is de kortste afstand tot het doel.

Opgave 13 (13.6.2012)

b. Eerst wordt A ontwikkeld (f-waarde $0+3=3$, met $f=g+h$) en komen B en C in de fringe, met waardes $2+0=2\rightarrow 3$ voor B ($2\rightarrow 3$ dankzij “pathmax”) en $1+2=3$ voor C.
Dan wordt B of C ontwikkeld. Stel eerst B. Dat levert een fringe met C, Z en X, met waardes 3, $6+0=6$ en $5+5=10$, respectievelijk. Dan wordt C ontwikkeld, wat knoop Z met waarde $4+0=4$ aan de fringe toevoegt. Nu wordt Z (met waarde 4) ontwikkeld: een doelknoop, dus het algoritme stopt. Kortste pad is A-C-Z met waarde 4.
Als eerst C en daarna B was ontwikkeld, ontstaat dezelfde fringe als zo-even. In feite lopen we door nevenstaande boom.



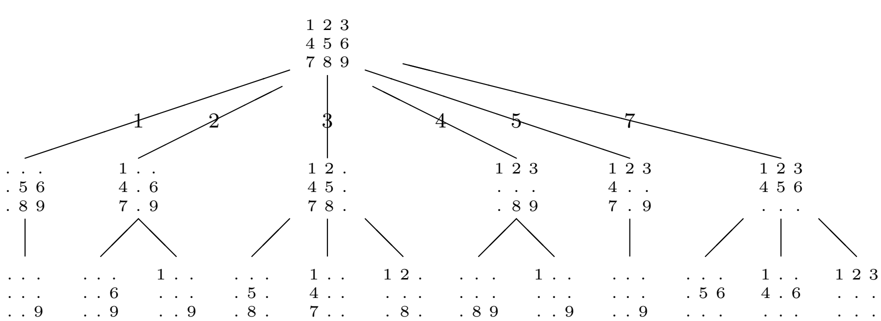
In dit geval, de graaf is ongericht, zouden we ook nog vanuit B weer terug naar A kunnen gaan, wat een kind $A(7)$ van B zou opleveren, etcetera. Laten we hopen dat we zo handig zijn dat niet te doen.

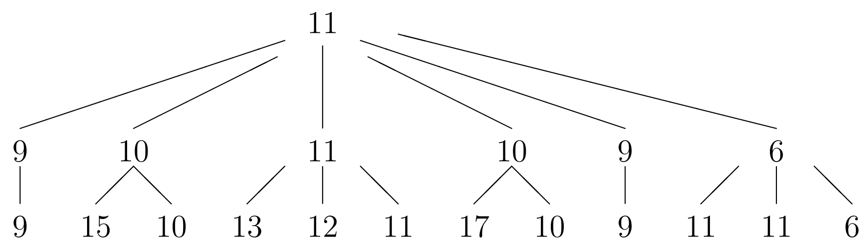
c. Bij IDA* beginnen we met een DFS (Depth First Search) met f-waarde van te expanderen knopen naar boven begrensd door de f-waarde van A, namelijk $f_{\text{lim}}=0+3=3$. We maken/bezoeken dan alle knopen van bovenstaande boom, in DFS, dus in volgorde A-B-X-Z-C-Z (of A-C-Z-B-X-Z) met nieuwe bovengrens op f gelijk aan $f_{\text{lim}}=4$. Daarna nog een zelfde DFS-wandeling, waarbij het algoritme stopt op het punt dat $Z(4)$ ge-expandeerd gaat worden.

d. De toestanden zijn nu: A, A', B, B', C, C', X', Z en Z' (waarbij een $'$ aangeeft dat we het pakje hebben). Er valt over te twisten of ook X, zonder pakje dus, bestaat. De beste h, die altijd exact de kortste afstand aangeeft, wordt: $A(12)$, $A'(4)$, $B(10)$, $B'(4)$, $C(13)$, $C'(3)$, $X'(7)$, $Z(14)$ en $Z'(0)$.
In het gunstigste geval loop je met IDA* in één keer van A naar B naar X naar B naar Z, waarbij alle f-waardes onderweg gelijk zijn aan 12. Kortom: één DFS, met $f_{\text{lim}}=12$.

Opgave 14 (4.6.2004)

a. De wortel heeft 6 kinderen en 12 kleinkinderen. Je krijgt bladeren 9 (enig kind, na eerste keus ’1’), 15, 10 (2 broers, na eerste keus ’2’), 13, 12, 11 (3 broers, na eerste keus ’3’), 17, 10 (2 broers, na eerste keus ’4’), 9 (enig kind, na eerste keus ’5’), 11, 11 en 6 (3 broers, na eerste keus ’7’). Zie onder.

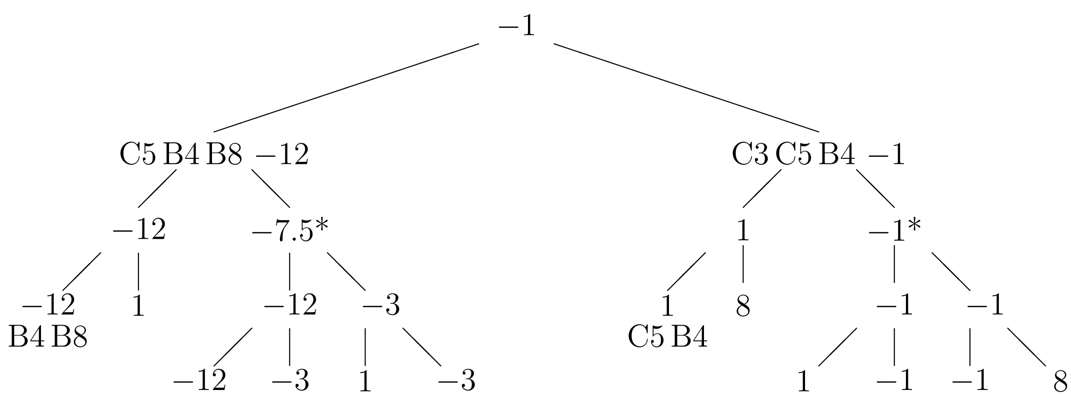




c. Voor **A** hoog: 11, namelijk het maximum van 9, 10, 11, 10, 9 en 6. Voor **A** laag: 9, namelijk het minimum van 9, 15, 13, 17, 9 en 11.

d. Maximaal 4 "geprunede" bladeren voor **A** hoog. Zorg dat het meest linker kind alvast 11 oplevert (dat is een knoop met 3 kinderen). Dan kunnen vervolgens alle rechter broers op het laagste nivo "gepruned" worden. Even opletten met de volgorde van de kinderen bij die andere knopen . . .
Maximaal 6 "geprunede" bladeren voor **A** laag. Zorg dat het meest linker kind alvast 9 oplevert. Dan kunnen vervolgens alle rechter broers "gepruned" worden. Bij de knoop met kinderen 11, 11 en 6 moet dan natuurlijk niet eerst de 6 bekeken worden. De boom die bij **a** is getekend, voldoet.

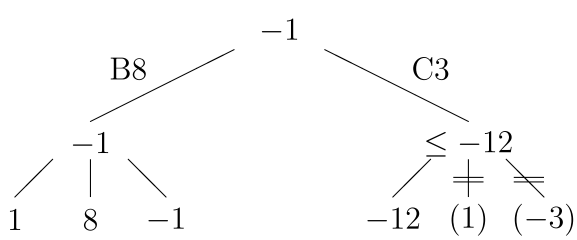
Opgave 16 (22.6.2011)
b.



Waardes waar Cindy wint zijn positief, voor Bert negatief. De twee knopen met een * zijn dobbelknopen; daar komt het gemiddelde van de kinderen. De wortel is een MAX-knoop, de andere knopen zijn MIN-knopen. De volgorde van de kinderen kan anders zijn. Bij twee bladeren staan de tegels — en dat zou bij alle knopen wel mogen.

c. Als je C5 $\uparrow$ C7 of B8 $\downarrow$ B6 doet, verandert de waarde in de wortel niet. Als je C3 $\uparrow$ C5 of B4 $\downarrow$ B2 doet, krijgt de wortel 1. En dat vindt Cindy fijn.

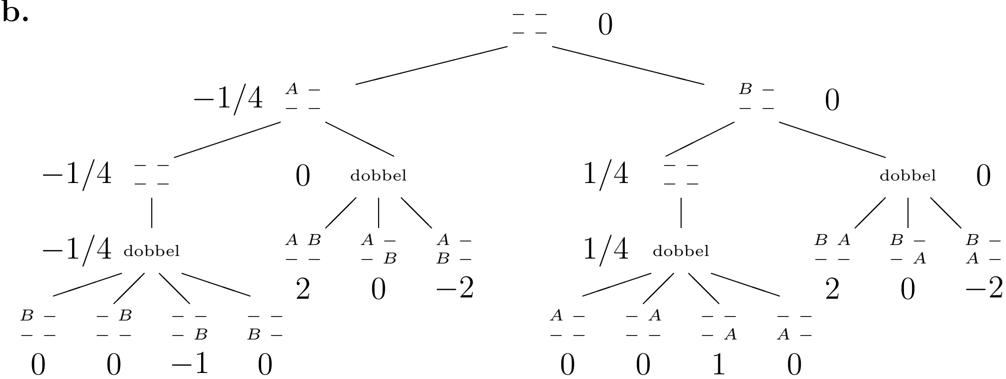
d.



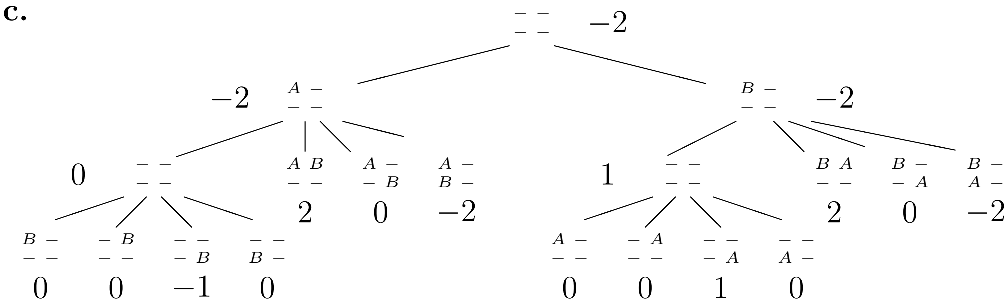
De twee gemarkeerde takken worden gepruned. Als je daar eerst als eerste kind -3 hebt (in plaats van -12), gaat het op dezelfde manier.

Opgave 17 (13.6.2012)

b.



c.



Opgave 19 (4.6.2006)

e. Gebruik de "most constraining variable" heuristiek voor de keus van de eerste te kleuren knoop: kies C, want die heeft de meeste buren: 4 stuks. Kies dan bijvoorbeeld A. Nu hebben B en D nog maar één mogelijke kleur, en dus kleuren we die nu (volgens de "minimum remaining values" heuristiek). Kies bijvoorbeeld vervolgens knoop E. Volgens de "least constraining value" heuristiek geven we die dezelfde kleur als C. De rest is eenvoudig te kleuren. (Er zijn andere oplossingen mogelijk.)

Opgave 23 (13.8.2001)

a. Een bias-knoop modelleert de drempelwaarde van de knoop waar hij naartoe wijst.

b. In formule, voor het gewicht $W_{j,i}$ van verborgen knoop j naar uitvoerknoop i:

$$\frac{\partial}{\partial W_{j,i}} \; 1/2 \; \sum_{\ell} \text{Error}_{\ell}^2 = -\text{Error}_i \, a_j \, g'(\text{invoer}_i) \qquad W_{j,i} \leftarrow W_{j,i} + \alpha \, \text{Error}_i \, a_j \, g'(\text{invoer}_i)$$

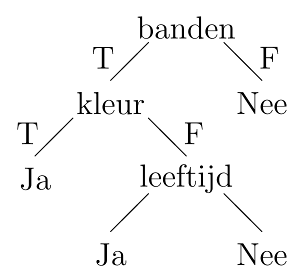
want $\text{Error}_i = T_i - O_i = T_i - g(\text{invoer}_i)$ en $\text{invoer}_i = \sum_{\ell} W_{\ell,i} a_{\ell}$; hierbij: α is de leersnelheid, T_i is de i-de target, a_{ℓ} de activatie van neuron ℓ, en g de activatie-functie. De gradiënt wijst in de richting van de grootste fout-stijging, dus er komt nog een extra $-$ bij als je een dalende fout wilt.

c. Bij vergelijkbare prestatie is —volgens Ockham— een eenvoudiger "verklaring" beter. Dus minder verborgen knopen graag, mits de prestatie er niet te veel op achteruit gaat.

d. Als de uitvoer 4 knopen voor de 4 windrichtingen heeft ("gedistribueerd coderen") kan het gebeuren dat ze alle vier bijvoorbeeld 1 geven: wat betekent dat? Handiger wellicht één knoop die de windrichting met één getal codeert, bijvoorbeeld de hoek ("locaal coderen"). Let er wel op dat dan de fout "op de cirkel" gemeten wordt, dus 359 graden is dichtbij 0 graden.

Opgave 24 (4.6.2004)
a. Zie sheet 9 van Neurale netwerken: `link`.
b. De positieve en negatieve voorbeelden zijn niet lineair te scheiden, zie sheet 13.
c. Zie sheet 26, eerste helft (of zelfs sheet 16). Dit leidt (met g de activatie-functie, en in_i de invoer van de i-de uitvoerkonoop) tot $W_{ij} \leftarrow W_{ij} + \alpha(T_i - O_i)g'(\text{in}_i)a_j$.
d. Definieer de fout als $\text{Error} = \sum_i (T_i - O_i)^2 + \sum_{i,j} W_{ij}^2$. In de partiële afgeleide $\partial\text{Error}/\partial W_{ij}$ krijg je er dan ook bij $\ldots + 2W_{ij}$. En de updateregel wordt dan $W_{ij} \leftarrow W_{ij} + \alpha(T_i - O_i)g'(\text{in}_i)a_j - \beta W_{ij}$. Die laatste min komt erbij omdat de gradiënt wijst in de richting waarin de fout juist het meest toeneemt. En β geeft aan hoe sterk dit meeweegt. In iedere trainingsstap stuur je alle gewichten een beetje richting 0.
Vaak wordt overigens $W_{j,i}$ gebruikt voor het gewicht op de tak van j naar i, in plaats van W_{ij}.

Opgave 27 (20.6.2006)
a. Zie de sheets: `link`.
b. Voor elk van de drie mogelijke attributen (kleur, banden en leeftijd) rekenen we de entropiewinst uit. Vooraf is deze 1 (want er zijn 2 positieve en 2 negatieve gevallen).

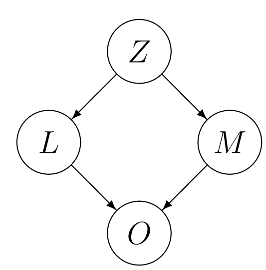


Voor kleur heb je voor True 1 positief en 1 negatief voorbeeld, dus entropie 1; voor False idem. Dus na afloop weer entropie 1, en dus Gain 0. Voor banden heb je voor True 2 positieve gevallen en 1 negatieve, dus entropie

$$-\frac{2}{3}\log_2\frac{2}{3} - \frac{1}{3}\log_2\frac{1}{3} \approx 0.9$$

Voor False 1 negatief geval: entropie 0. Dus na afloop $3/4 \cdot 0.9 + 1/4 \cdot 0 \approx 0.7$, en Gain $1 - 0.7 = 0.3$. Voor leeftijd komt hier hetzelfde uit, en het is beter dan Gain 0. Kies dus uit banden en leeftijd, zeg banden, in de wortel van de boom.
Met de gevallen 1, 3 en 4 gaan we door, waarbij we moeten kiezen uit de attributen kleur en leeftijd. Voor kleur geldt dat voor True we 1 positief geval hebben en voor False 1 positief en 1 negatief geval, dus vooraf 0.9 en achteraf $1/3 \cdot 0 + 2/3 \cdot 1 \approx 0.7$ met Gain $0.9 - 0.7 = 0.2$. Voor leeftijd blijkt de Gain heztelfde. Kies nu bijvoorbeeld kleur, waarna nog een vraag naar leeftijd nodig is.
c. Maar er is een boom van hoogte 2 mogelijk: vraag eerst naar kleur, en gebruik banden respectievelijk leeftijd om daarna direct goed te scheiden. ID3 is een gretig algoritme!

Opgave 30 (13.8.2001)
a. Als trouwens $P(o|\ell, m) = 0.75$ zou zijn, was O een "noisy-or" geweest (want ook $P(o|\overline{\ell}, \overline{m}) = 0$).
b. Er geldt overigens: **b** is diagnostisch en **c** causaal. We gaan eerst **c** doen, dat is eenvoudiger, en gebruiken dan $P(z|o) = P(o|z)P(z)/P(o)$ (Bayes). Nu is



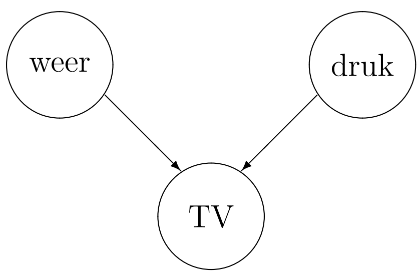
$$
\begin{aligned}
P(o|z) &= P(o|\ell,m)P(\ell|z)P(m|z) + P(o|\overline{\ell},m)P(\overline{\ell}|z)P(m|z) + \\
&\quad\; P(o|\ell,\overline{m})P(\ell|z)P(\overline{m}|z) + P(o|\overline{\ell},\overline{m})P(\overline{\ell}|z)P(\overline{m}|z) \\
&= 0.9 * 0.6 * 0.8 + 0.5 * 0.4 * 0.8 + 0.5 * 0.6 * 0.2 + 0.0 * 0.4 * 0.2 = 0.652.
\end{aligned}
$$

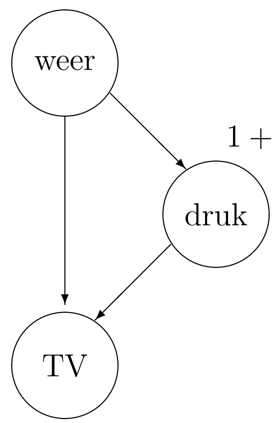
Analoog $P(o|\overline{z}) = \ldots = 0.197$, en dus $P(o) = P(o|z)P(z) + P(o|\overline{z})P(\overline{z}) = 0.334$, waaruit volgt: $P(z|o) = 0.652 * 0.3/0.334 \approx 0.59$.

En dan toch maar **b**: $P(z|\ell,m) = P(\ell,m|z)P(z)/P(\ell,m)$ (met Bayes). Dan moeten we ook nog doen: $P(\ell,m) = P(\ell,m,z)+P(\ell,m,\overline{z}) = P(\ell|z)P(m|z)P(z)+P(\ell|\overline{z})P(m|\overline{z})P(\overline{z})$. En tot slot: $P(\ell,m|z) = P(\ell|z)P(m|z)$.

Opgave 31 (5.6.2002)

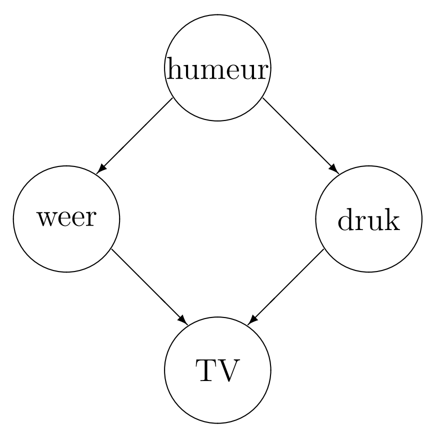
ambtenaar
$1+1+4=6$ getallen



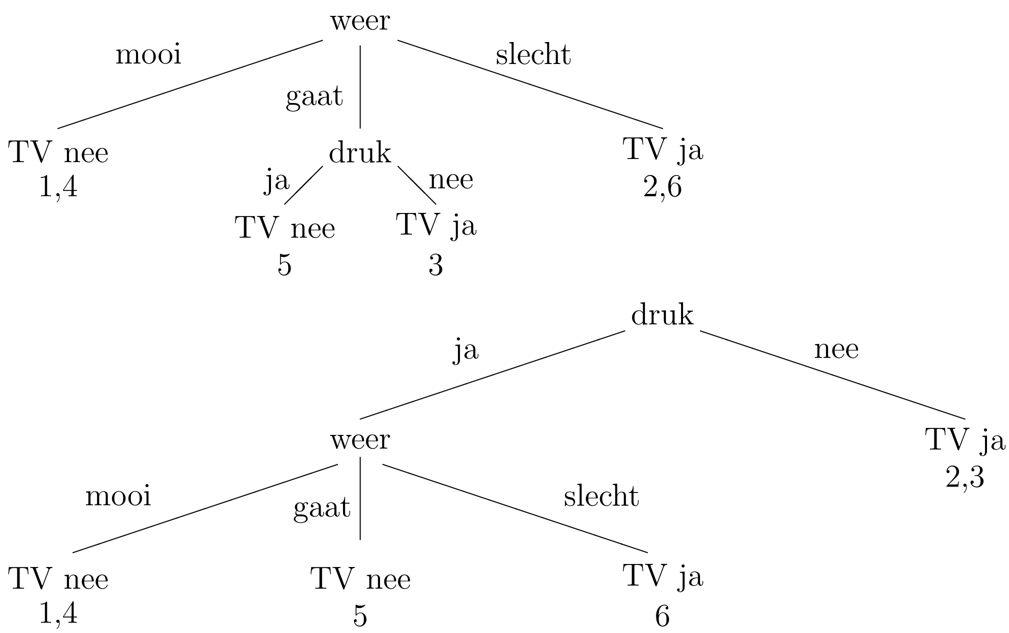


werker in
strandtent
$1+2+4=7$ getallen

Over de tak van weer naar TV (strandtent) valt wellicht te discussiëren.
c. Diagnostisch: P(mooi weer | TV ja), causaal: P(TV nee | mooi weer), intercausaal: P(mooi weer | TV ja $\wedge$ wel druk) (alle drie voor ambtenaar), mixed: P(niet druk | TV ja $\wedge$ slecht weer) (voor werker in strandtent).
e. Met "cutset" humeur: maak twee netwerken, één voor humeur goed en één voor humeur slecht — zonder kringen.



Opgave 32 (5.6.2002)
a.



b. $I(3/6, 3/6) = 1$, de entropie vooraf. Na de eerste weer-vraag:

$$1/3 \times I(1,0) + 1/3 \times I(1/2, 1/2) + 1/3 \times I(0,1) = 1/3 \times 0 + 1/3 \times 1 + 1/3 \times 0 = 1/3,$$

met gain: $1 - 1/3 = 2/3$.

Opgave 33 (25.6.2003)
a. $P(w)$, $P(a|w)$, $P(a|\overline{w})$, $P(e|w,a)$, $P(e|\overline{w},a)$, $P(e|w,\overline{a})$, $P(e|\overline{w},\overline{a})$, $P(b|a)$, $P(b|\overline{a})$, $P(k|a,e)$, $P(k|\overline{a},e)$, $P(k|a,\overline{e})$ en $P(k|\overline{a},\overline{e})$
b. $P(B|A) = P(B|A,W)$
c. $P(b|\overline{w}) = P(b|a)P(a|\overline{w}) + P(b|\overline{a})P(\overline{a}|\overline{w})$ (en $P(\overline{a}|\overline{w}) = 1 - P(a|\overline{w})$).
d. $P(k|\overline{w}) = P(k|a,e)P(e|\overline{w},a)P(a|\overline{w}) + P(k|\overline{a},e)P(e|\overline{w},\overline{a})P(\overline{a}|\overline{w})$
$+ P(k|a,\overline{e})P(\overline{e}|\overline{w},a)P(a|\overline{w}) + P(k|\overline{a},\overline{e})P(\overline{e}|\overline{w},\overline{a})P(\overline{a}|\overline{w})$
e. $P(a|\overline{b},w) = P(\overline{b}|a,w)P(a|w)/P(\overline{b}|w) = P(\overline{b}|a)P(a|w)/P(\overline{b}|w)$ (via Bayes, **b** en **c**)
f. **c** en **d** zijn causaal, **e** mixed.

Opgave 34 (3.6.2005)
b. $P(E|F,N,W) = P(E|F,N)$
c. $P(e|\overline{w}) = P(e|f,n)P(n|f,\overline{w})P(f|\overline{w}) + P(e|f,\overline{n})P(\overline{n}|f,\overline{w})P(f|\overline{w}) +$
$P(e|\overline{f},n)P(n|\overline{f},\overline{w})P(\overline{f}|\overline{w}) + P(e|\overline{f},\overline{n})P(\overline{n}|\overline{f},\overline{w})P(\overline{f}|\overline{w})$; causaal
d. $P(f|\overline{e},w) = P(\overline{e}|f,w)P(f|w)/P(\overline{e}|w)$, waarbij $P(f|w)$ en $P(\overline{e}|w)$ (zie **c**) bekend zijn, en $P(\overline{e}|f,w) = P(\overline{e}|f,n)P(n|f,w) + P(\overline{e}|f,\overline{n})P(\overline{n}|f,w)$; mixed

Opgave 39 (24.6.2009)
Zie uitwerking tentamen: `link`.

Opgave 40 (24.6.2009)
Zie uitwerking tentamen: link.

Opgave 41 (24.6.2009)
Zie uitwerking tentamen: link.

Opgave 42
e. In totaal vijf stuks: $P(a)$, $P(b \mid a)$, $P(b \mid \overline{a})$, $P(c \mid b)$ en $P(c \mid \overline{b})$.
h. Soms zijn er meerdere doelen van waaruit moet worden terug-geredeneerd. En soms is het bepalen van voorgangers lastig. Beide zijn het geval bij schaken.
j. Er zijn voor een zaal in totaal vier mogelijkheden: 'niks', 'beamer', 'computer' en 'beamer&computer'. We weten dat er in zaal X veel stroom wordt verbruikt. De twee modellen die hiermee overeenkomen zijn 'computer' en 'beamer&computer'. Deze impliceren "Er staat in zaal X een computer".
Zie sheet 11 bij de Wumpus wereld: link.

Zie verder de sheets en het boek.