

# Opgaven Kunstmatige intelligentie — 4

## maandag 2 mei 2011

**Opgave 28.** (opgave tentamen 12 augustus 2002)

Stel dat we een handelsreizigersprobleem op willen lossen, en dat we dat met een *genetisch algoritme* willen doen. Gegeven zijn steden en al hun onderlinge afstanden, gezocht een kortste route die alle steden bezoekt, en eindigt waar de route begonnen is.

- Geef (in woorden) aan hoe een genetisch algoritme er in het algemeen uit ziet.
- Geef (in woorden) aan hoe een *evaluatiefunctie* (= *fitnessfunctie*) er in dit speciale geval uit zou kunnen zien.
- Geef aan hoe *crossover* en *mutatie* in dit speciale geval zouden kunnen werken.
- Wanneer heet een genetisch algoritme *elitair*?

**Opgave 29.** (opgave tentamen 20 juni 2006)

We hebben een database waarin een aantal voorbeelden aan de hand van een stel attributen beschreven worden; verder hebben we een Ja/Nee-classificatie voor deze voorbeelden.

- Leg uit hoe het *ID3-algoritme* werkt. Geef hierbij de definitie van de entropie (uitgedrukt in  $p$ , het aantal positieve voorbeelden en  $n$ , het aantal negatieve). Geef met name aan wat er moet (kan) gebeuren in de vier verschillende “standaard”-gevallen, bijvoorbeeld als er bij het splitsen geen voorbeelden meer zijn.
- De volgende database geeft situaties uit het verleden aan. We willen een beslissingsboom maken om te kijken of een fiets goed is via drie Boolese attributen:

nr	kleur	banden	leeftijd	fiets goed?
1	T	T	F	Ja
2	T	F	F	Nee
3	F	T	T	Ja
4	F	T	F	Nee

Voer het ID3-algoritme uit, met uitleg (kies random in geval van gelijke kwaliteit).

- Is er een correcte beslissingsboom met een kleinere diepte? Zo ja, geef deze; zo nee, waarom niet?

**Opgave 30.** (opgave tentamen 19 juni 2007)

We hebben een database waarin een aantal voorbeelden aan de hand van een stel attributen beschreven worden; verder hebben we een Ja/Nee-classificatie voor deze voorbeelden.

- Leg uit hoe het *ID3-algoritme* werkt. Geef hierbij de definitie van de entropie (uitgedrukt in  $p$ , het aantal positieve voorbeelden en  $n$ , het aantal negatieve). Geef met name aan wat er moet (kan) gebeuren in de vier verschillende “standaard”-gevallen, bijvoorbeeld als er bij het splitsen geen voorbeelden meer zijn.

model	geheugen	processor	rest	computer OK?
1	T	T	F	Ja
2	T	F	F	Nee
3	F	T	T	Ja
4	F	T	F	Nee

**b.** De database geeft situaties uit het verleden aan. We willen een beslissingsboom maken om te kijken of een computer voldoet via drie Boolese attributen.

Voer het ID3-algoritme uit, met uitleg (kies random in geval van gelijke kwaliteit).

**c.** Is er een correcte beslissingsboom met kleinere hoogte?

**d.** Stel dat we dit zelfde probleem met een *neuraal netwerk* met één verborgen laag met 3 knopen willen oplossen; het netwerk heeft ook bias-knopen. Teken dit netwerk.

**Opgave 31.** (opgave tentamen 12 augustus 2002)

We hebben een database waarin een aantal voorbeelden aan de hand van een stel attributen beschreven worden; verder hebben we een Ja/Nee-classificatie voor deze voorbeelden.

**a.** Leg uit hoe het *ID3-algoritme* werkt. Geef hierbij de definitie van de entropie  $I(p/(n+p), n/(n+p))$  ( $p$  is het aantal positieve voorbeelden,  $n$  het aantal negatieve). Geef met name aan wat er moet (kan) gebeuren in de vier verschillende “standaard”-gevallen, bijvoorbeeld als er bij het splitsen geen voorbeelden meer zijn.

**b.** De volgende database geeft situaties uit het verleden aan. We willen een beslissingsboom maken om te kijken of een film goed of slecht is, op grond van de acteurs en de lengte. Maak de twee mogelijke beslissingsbomen met hoogte 2.

dag	acteurs	lengte	film goed?
1	slecht	kort	ja
2	goed	lang	ja
3	gaat	lang	ja
4	slecht	kort	nee
5	gaat	kort	nee
6	goed	kort	ja

**c.** Reken voor de boom met als wortelvraag de vraag naar de acteurs, de entropie-winst (“gain”) uit voor deze eerste vraag.

**Opgave 32.** (opgave tentamen 13 augustus 2001)

Lex en Mary gaan soms uit, en komen elkaar dan wel eens tegen. Of ze uitgaan hangt vooral van het weer af. We hebben de volgende kansen, waarbij  $Z$  staat voor zonnig weer,  $L$  respectievelijk  $M$  voor Lex/Mary gaat uit, en  $O$  voor ontmoeting; een streep boven een letter staat voor de ontkenning; zo betekent  $\bar{Z}$  niet zonnig. Er geldt:  $P(Z) = 0.3$ ,  $P(L|Z) = 0.6$ ,  $P(L|\bar{Z}) = 0.1$ ,  $P(M|Z) = 0.8$ ,  $P(M|\bar{Z}) = 0.3$ ,  $P(O|L, M) = 0.9$ ,  $P(O|\bar{L}, M) = 0.5$ ,  $P(O|L, \bar{M}) = 0.5$  en  $P(O|\bar{L}, \bar{M}) = 0.0$ .

**a.** Teken het bijbehorende Bayesiaanse netwerk.

**b.** Wat is de kans dat het zonnig is gegeven dat Lex en Mary elkaar ontmoeten? (Geef zonder al te veel details aan hoe de berekening verloopt.)

**c.** En wat is de kans dat Lex en Mary elkaar ontmoeten als het zonnig is?

**d.** Er zijn vier soorten queries; noem deze en leg ze kort uit. Tot welke soort horen de queries van **b** en **c**?

**e.** Maak een nieuw netwerk, waarbij de twee knopen die met uitgaan te maken hebben gecombineerd worden tot één knoop. Geef de bijbehorende voorwaardelijke kanstabellen. Wat is het voordeel van een dergelijke aanpassing van het netwerk?

**Opgave 33.** (opgave tentamen 5 juni 2002)

**a.** Of we gaan TV-kijken hangt af van het weer (mooi/slecht) en of we het druk hebben (ja/nee). Teken het bijbehorende *Bayesiaanse netwerk*, zowel voor iemand die in een

strandtent werkt als voor een ambtenaar. Zorg voor twee verschillende netwerken, en maak de keuzes aannemelijk. Welke kanstabellen moeten gegeven worden?

**b.** Er zijn vier soorten queries; noem deze en leg ze kort uit.

**c.** Geef van elke soort query een voorbeeld uit een van de netwerken van **a**.

**d.** In een verre toekomst hangen voor een ambtenaar zowel het weer als het feit dat we het druk hebben af van het humeur van de leider. Maak allereerst het bijbehorende Bayesiaanse netwerk. Leg uit wat de “cutset conditioning methode” is, hoe deze werkt aan de hand van dit netwerk, en waarvoor deze in het algemeen gebruikt wordt.

**Opgave 34.** (opgave tentamen 5 juni 2002)

We hebben een database waarin een aantal voorbeelden aan de hand van een stel attributen beschreven worden; verder hebben we een Ja/Nee-classificatie voor deze voorbeelden.

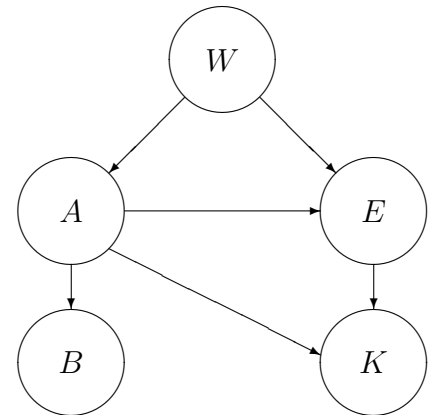
**a.** De volgende database geeft situaties uit het verleden aan. We willen een beslissingsboom maken om te kijken of we al of niet gaan TV-kijken, op grond van weer (mooi/gaat/slecht) en drukte. Maak de twee mogelijke beslissingsbomen met hoogte 2.

dag	weer	druk	TV-kijken
1	mooi	ja	nee
2	slecht	nee	ja
3	gaat	nee	ja
4	mooi	ja	nee
5	gaat	ja	nee
6	slecht	ja	ja

**b.** Reken voor de boom met als wortelvraag de vraag naar het weer, de entropie-winst (ID3, “gain”) uit voor deze eerste vraag.

**Opgave 35.** (opgave tentamen 25 juni 2003)

**a.** We hebben een *Bayesiaans netwerk* voor het verband tussen het winnen van een prijs ( $W$ ) en het blij zijn van een gezin, bestaande uit ouders Alfred ( $A$ ) en Ellen ( $E$ ), zoon Kurt ( $K$ ) en hond Bonzo ( $B$ ). We gebruiken kleine letters als volgt:  $b$  betekent dat Bonzo blij is,  $\bar{w}$  betekent dat de prijs niet gewonnen is, etcetera. Welke kanstabellen moeten gegeven zijn, en met welke kansen erin?



**b.** Geef een voorbeeld van een voorwaardelijke onafhankelijkheid, beginnend met  $P(B|A) = \dots$

**c.** Druk de kans  $P(b|\bar{w})$  dat Bonzo blij is, gegeven dat de prijs niet is gewonnen, uit in uit de tabellen bekende kansen.

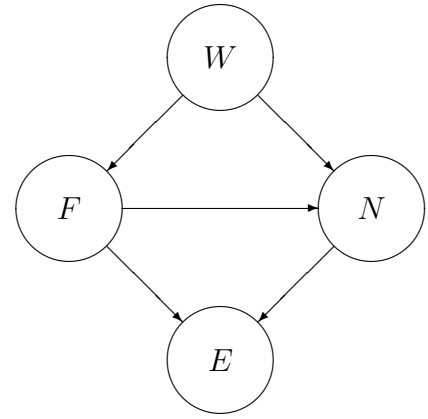
**d.** Idem voor Kurt:  $P(k|\bar{w})$ .

**e.** Druk de kans  $P(a|\bar{b}, w)$  dat Albert blij is, gegeven dat de prijs is gewonnen, maar Bonzo niet blij is, uit in uit de tabellen bekende kansen.

**f.** Er zijn vier soorten queries; noem deze en leg ze kort uit. Tot welke soort(en) behoren de queries van **c**, **d** en **e**?

**Opgave 36.** (opgave tentamen 3 juni 2005)

**a.** We hebben een *Bayesiaans netwerk* voor het verband tussen het mooie weer ( $W$ ), de verkiezingsuitslag ja/nee in Frankrijk ( $F$ ) en Nederland ( $N$ ), en de totale Europese uitslag ( $E$ ). We gebruiken kleine letters als volgt:  $e$  betekent dat de Europese uitslag ja (true) is,  $\bar{w}$  betekent dat het weer slecht (false) is, etcetera. Welke kanstabellen moeten gegeven zijn, en met welke kansen erin?



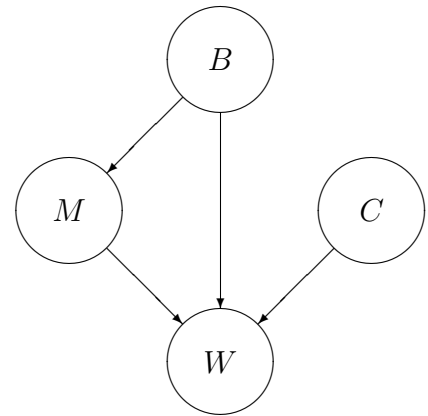
**b.** Geef een voorbeeld van een *voorwaardelijke onafhankelijkheid*:  $P(E|\dots) = \dots$

**c.** Druk de kans  $P(e|\bar{w})$  dat Europa ja zegt, gegeven dat het weer slecht is, uit in uit de kanstabellen van **a** bekende kansen.

**d.** Druk de kans  $P(f|\bar{e}, w)$  dat Frankrijk ja stemt, gegeven dat Europa nee zegt, terwijl het mooi weer is, uit in uit de kanstabellen van **a** bekende kansen.

**Opgave 37.** (opgave tentamen 21 augustus 2006)

**a.** We hebben een *Bayesiaans netwerk* voor het verband tussen de motivatie ( $M$ ), de conditie ( $C$ ), de beloning ( $B$ ), en de winst bij een sportwedstrijd ( $W$ ). We gebruiken kleine letters als volgt:  $b$  betekent dat de beloning goed (true) is,  $\bar{c}$  betekent dat de conditie slecht (false) is, etcetera. Welke kanstabellen moeten gegeven zijn, en met welke kansen erin?



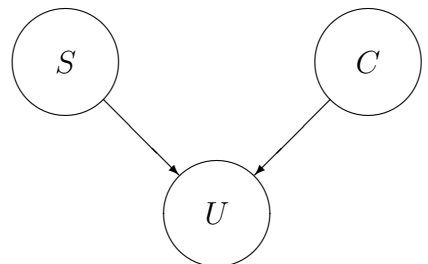
**b.** Hoeveel kansen zijn dit maximaal bij een Bayesiaans netwerk met 4 knopen, en vul het netwerk aan tot zo'n netwerk.

**c.** Druk voor het oorspronkelijke netwerk de kans  $P(\bar{w}|b)$  dat de winst niet behaald wordt, gegeven dat de beloning goed is, uit in uit de kanstabellen van **a** bekende kansen.

**d.** Druk de kans  $P(m|w, \bar{b})$  dat de motivatie goed is, gegeven dat de winst behaald wordt, bij slechte beloning, uit in uit de kanstabellen van **a** bekende kansen.

**Opgave 38.** (opgave tentamen 25 juni 2008)

We hebben een *Bayesiaans netwerk* waarin blijkt dat een goede uitslag van de wedstrijd ( $U$ ) afhangt van de goede conditie ( $C$ ) en het goede salaris ( $S$ ). We gebruiken kleine letters als volgt:  $u$  betekent dat de uitslag goed (true) is,  $\bar{c}$  betekent dat we slechte conditie hebben (false), etcetera. Neem aan dat  $U$  een *noisy-or* knoop is, dat  $P(s) = 0.4$ ,  $P(c) = 0.4$ ,  $P(u|c, \bar{s}) = 0.7$  en  $P(u|\bar{c}, s) = 0.8$ .



**a.** Geef de volledige kanstabellen die bij het netwerk horen.

**b.** (4 punten) Bereken  $P(\bar{s}|u, c)$ .

**c.** We breiden het netwerk uit met een knoop  $T$ , voor een goede trainer, die rechtstreeks de conditie beïnvloedt, en niet het salaris:  $P(c|t) = 0.6$  en  $P(c|\bar{t}) = 0.1$ , terwijl  $P(t) = 0.6$ . Zijn deze waarden consistent met de bij **a** gegeven waarden?

**d.** Druk de kans  $P(t|c, \bar{u})$  uit in bekende kansen.

- e. Er zijn vier soorten queries; noem deze en leg ze kort uit. Tot welke soort(en) behoren de queries van **b** en **d**?
- f. Stel dat de trainer ook de salarissen bepaalt. Teken het nieuwe netwerk, en leg kort een methode uit om queries effectief te kunnen behandelen.

**Opgave 39.** (opgave tentamen 25 juni 2008)

We hebben een database waarin een aantal voorbeelden aan de hand van een stel attributen beschreven worden; verder hebben we een Ja/Nee-classificatie voor deze voorbeelden.

a. Leg uit hoe het *ID3-algoritme* werkt. Geef hierbij de definitie van de entropie (uitgedrukt in  $p$ , het aantal positieve voorbeelden en  $n$ , het aantal negatieve). Geef met name aan wat er moet (kan) gebeuren in de vier verschillende “standaard”-gevallen, bijvoorbeeld als er bij het splitsen geen voorbeelden meer zijn.

b. De volgende database geeft situaties uit het verleden aan. We maken een beslissingsboom om te kijken of een voetballer goed is via drie Boolese attributen:

speler	koppen	scoren	verdedigen	goed?
1	T	T	F	Ja
2	T	F	F	Nee
3	F	F	T	Ja
4	F	T	F	Nee
5	F	T	T	Nee

Voer het ID3-algoritme uit, met uitleg (kies zelf in geval van gelijke kwaliteit).

c. Beschrijf kort het leren van een Boolese functie voor dit probleem, met behulp van een *current best hypothesis*. Leg in ieder geval *false positives/negatives* uit.

d. Pas deze methode toe, beginnend met hypothese “koppen”.

**Opgave 40.** (opgave tentamen 24 juni 2009)

We bekijken de *4-puzzel*. Er zijn 3 tegels, A, B en C, en een lege plek aangeduid met -. Een zet (kosten zijn steeds 1) bestaat uit het verschuiven van een tegel naar de lege plek, mits deze horizontaal of verticaal aan elkaar grenzen. De begintoestand (links) en eindtoestand (rechts) staan hierboven.

- A    B -  
C B    A C

a. Geef de gehele toestand-actie-ruimte/graaf. Hint: er zijn 12 toestanden.

b. Als admissibele heuristische functie  $h$  nemen we de functie die zegt hoeveel tegels er niet op de juiste plek liggen (voor de begintoestand zijn dat er 3). Voer het IDA\*-algoritme uit voor deze graaf. Geef duidelijk aan hoe het algoritme verloopt, en met name in welke volgorde knopen ontwikkeld worden. Indien er keuze mogelijk is in deze volgorde, geef ze dan allemaal.

c. Idem, maar nu met als heuristische functie de functie die voor alle tegels uitrekent hoeveel plekken ze minstens moeten verschuiven, en deze waarden optelt (voor de begintoestand: 5).

d. Stel we voegen een extra tak toe tussen de begintoestand en een van de directe burenen van de doeltoestand. Wat zou er nu mis gaan bij onderdeel **d**, en waarom?

**Opgave 41.** (opgave tentamen 24 juni 2009)

Max en Mini spelen een tweepersoons spel. De beginspeler, Max, mag kiezen voor versie *Links* of versie *Rechts*; dit is het enige wat Max doet. Daarna is Mini aan de beurt. Voor de gebruikte munt geldt: kans  $1/3$  om “kop” te gooien, en kans  $2/3$  op “munt”. De spelers gebruiken vier getallen:  $A = 2$ ,  $B = 5$ ,  $C = 3$  en  $D = 7$ . Uiteraard streeft Max naar

een zo hoog mogelijke uitkomst, en Mini naar een zo laag mogelijke. We spelen straks verschillende varianten.

a. Beschrijf in woorden het *expecti-minimax-algoritme*.

b. In de versie *Links* wordt eerst een munt gegooid, en bij “kop” kiest Mini uit  $B$  en  $D$ , en bij “munt” uit  $A$  en  $C$ . In de versie *Rechts* wordt eerst een munt gegooid, en bij “kop” kiest Mini uit  $A$  en  $B$ , en bij “munt” uit  $C$  en  $D$ . Maak de spelboom, en bereken de expecti-minimax-waarde.

c. Iemand beweert het volgende: “We vervangen de waarde van  $B$  door 50, die van  $C$  door 30 en die van  $D$  door 70. Er blijft gelden dat  $A < C < B < D$ . Dus het maakt voor de analyse van het spel bij **b** niet uit.” Is dit juist of niet, en waarom?

d. Nu mag Mini voordat Max bij **b** de versie kiest één van de getallen  $A$ ,  $B$ ,  $C$  of  $D$  met 1 verlagen. Doe een goede suggestie.

**Opgave 42.** (opgave tentamen 24 juni 2009)

Gegeven zijn 2 negatieve voorbeelden:  $P = (6, 0)$  en  $Q = (3, 2)$ ; en 1 positief voorbeeld  $R = (1, 0)$ . We zoeken een *perceptron* met 2 invoeren en 1 uitvoer om deze 3 voorbeelden goed te classificeren: voor een negatief voorbeeld moet 0 worden opgeleverd, voor een positief voorbeeld 1.

a. (5 punten) Geef de benodigde waarden voor de gewichten en de drempelwaarde. Teken ook een plaatje van het netwerk.

b. (5 punten) Idem, met positief voorbeeld  $S = (4, 0)$  erbij.

c. (5 punten) Idem, maar vervang  $S = (4, 0)$  door  $S = (5, 2)$ .

d. (5 punten) Wat heeft *Ockham's razor* met neurale netwerken te maken?

**Opgave 43.**

a. Leg uit wat het verschil is tussen “steady-state” en “generationele” strategieën bij Genetische Algoritmen.

b. Bij een *Genetisch Algoritme* kun je *uniform crossover* en *single-point crossover* gebruiken. Wat is het verschil?

c. We willen met het ID3-algoritme een beslissingsboom maken voor een database met situaties uit het verleden, waarbij we de waarden van diverse attributen weten. Stel er is een attribuut bij met veel verschillende waarden. Wat kan er fout gaan bij het ID3-algoritme, en wat kun je hier tegen doen?

d. Stel je bent met een “current best hypothesis” aan het leren. Wat kun je doen als je een “false positive” voorbeeld tegenkomt, en wat bij een “false negative” voorbeeld? Geef uitleg.

e. Hoeveel en welke kansen moeten er gegeven worden bij het volgende *Bayesiaanse netwerk* met drie knopen  $A$ ,  $B$  en  $C$ :  $A \longrightarrow B \longrightarrow C$ ?

f. Iemand beweert (ten onrechte, overigens) dat schaken *ultra-zwak* is opgelost. Wat bedoelt hij/zij met deze opmerking?

g. Formuleer het volgende probleem formeel als een *Constraint Satisfaction Problem*: We willen nevenstaande graaf zodanig met  $\leq 3$  kleuren inkleuren dat maximaal één paar aangrenzende knopen dezelfde kleur heeft.

