

Tentamen Kunstmatige intelligentie

Universiteit Leiden — Informatica

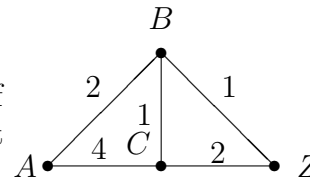
Maandag 29 mei 2017, 10:00–13:00 uur



Geef korte, duidelijke toelichting. Cijfers: www.liacs.leidenuniv.nl/~kosterswa/AI/res.html.

Opgave 1: A^*/IDA^* (20 punten)

a. (6 punten) Leg het A^* -*algoritme* en het IDA^* -*algoritme* uit. Geef de verschillen duidelijk aan. Geef expliciet de formule voor f (wat stellen g en h voor?) en denk aan de stop-conditie.



- b. (3 punten) We bekijken bovenstaande *ongerichte* graaf. We beginnen in A , en het doel is om zo snel mogelijk weer in A te komen, na Z bezocht te hebben. Een tak mag in een route hooguit één maal gebruikt worden. Een toestand wordt vastgelegd door de locatie = knoop en het reeds afgelegde pad. Als heuristiek h bepalen we in een toestand de lengte van een kortste route naar A , weer zonder dubbele takken, via Z (als Z nog niet bezocht was, anders rechtstreeks). We letten er hierbij niet op of eerder bezochte takken opnieuw voorkomen. Geef h voor de zeven mogelijke situaties, zoals “in B , nog niet in Z geweest”.
- c. (5 punten) We definiëren $h' = h$, waarbij de heuristische waarden die bij A en Z horen op 0 worden gezet. Voer het A^* -*algoritme* uit voor deze h' . Gebruik zonedig de *pathmax equation*. Geef duidelijk aan hoe het algoritme verloopt, en in welke volgorde de knopen ontwikkeld worden. Indien er hierbij keuzes mogelijk zijn, geef ze dan allemaal.
- d. (3 punten) Idem, maar nu het IDA^* -*algoritme* voor h' , in het gunstigste geval.
- e. (3 punten) Heuristiek h' is niet *consistent*. Waaruit blijkt dat? Leg het begrip ook uit.

Opgave 2: α - β -*algoritme* (20 punten)

Dick en Danny spelen een kaart/dobbelspel. Dick begint. Hij gooit een eerlijke driekantige dobbelsteen, die de rij bepaalt waaruit Danny een kaart kiest. De rij en kolom van deze kaart worden verwijderd

1:	♥4	♠2	◇8
2:	♥2	♣2	◇6
3:	♣4	◇2	♣5

uit het schema, en Dick kiest één van de vier overgebleven kaarten, waarbij hij een rode kaart (◇, ♥) moet kiezen als Danny een zwarte (♣, ♠) heeft gekozen, en andersom. Als Dick geen keuze heeft, verliest hij direct met 8 punten. De gekozen kaarten worden vergeleken, en de “hoogste” kleur wint: ♣ < ◇ < ♥ < ♠. In alle gevallen wint men met het absolute verschil in aantal punten (hoger is beter); en verschil 0 betekent remise.

Voorbeeld: Dick dobbelt 2, waarna Danny ♣2 kiest. Daarna kiest Dick ◇8 (hij had ook ♥4 kunnen kiezen). Dick wint met $|8 - 2| = 6$ punten: ♣ < ◇.

- a. (4 punten) Beschrijf in woorden het *expecti-minimax-algoritme*. Wat gebeurt er in max-knopen, min-knopen, kansknopen, bladeren (twee gevallen), ...?
- b. (7 punten) Maak de spelboom (met zeventien bladeren) en bereken expecti-minimax-waarde.
- c. (3 punten) Beide spelers krijgen het aanbod ♣5 door ♠5 te vervangen in het schema. Zou iemand van hen hierop in gaan, en waarom?
- d. (6 punten) Bekijk de spelboom voor het geval Dick, in plaats van te dobbelen, zelf de rij mag kiezen; geef tevens de negen kleinkinderen vast de waardes uit **b** (anders wordt de boom te groot). Voer het α - β -*algoritme* uit. Zorg ervoor dat de ordening van de knopen zo is dat er *maximaal* gesnoeid kan worden. Geef een korte rechtvaardiging voor het snoeien.

Opgave 3: ID3-algoritme & Neurale netwerken (20 punten)

We willen de kwaliteit (slecht/goed) van een oude computer inschatten, op grond van een vijftal eerdere observaties:

	chip	gewicht	scherm	kwaliteit
1	PENTIUM	++	mooi	g
2	PENTIUM	+	lelijk	g
3	80286	+	lelijk	s
4	8086	-	lelijk	s
5	80286	+	mooi	g

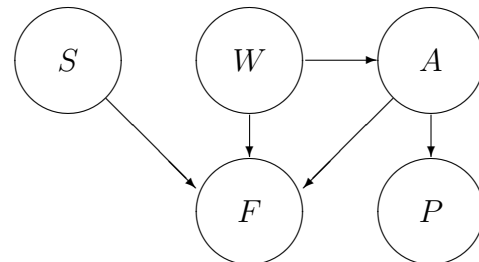
a. (7 punten) Leg het ID3-*algoritme* uit. Onderscheid hierbij vier gevallen voor de knopen: geen voorbeelden meer, ... Geef de benodigde formules, waaronder die voor de entropie $H(p, n)$ in geval van p positieve en n negatieve gevallen; en die voor de entropie na afloop van het splitsen op een attribuut met ν waardes.

b. (8 punten) Voer het ID3-algoritme uit voor bovenstaande database. Doel is **kwaliteit**. Gebruik hierbij $-\frac{2}{3} \log_2 \frac{2}{3} - \frac{1}{3} \log_2 \frac{1}{3} \approx 0.9$.

c. (5 punten) We willen vervolgens een *Neuraal netwerk* maken dat met behulp van **chip**, **scherm** en **kwaliteit** het drie-waardige attribuut **gewicht** voorspelt. Dit laatste kan op twee manieren in de uitvoer van het netwerk worden gecodeerd. Hoe heten de mogelijkheden, en wat is het verschil?

Opgave 4: Bayesiaanse netwerken (20 punten)

We hebben een *Bayesiaans netwerk* met vijf knopen: W (weer), F (file), S (spitsuur), A (aanrijding) en P (politie), zie hiernaast.



a. (3 punten) Hoeveel en welke kansen moeten hierbij gegeven zijn?

b. (4 punten) Hoeveel kansen zijn dit maximaal als elk van de vijf knopen hoogstens twee inkomende pijlen zou hebben? Teken een bijbehorend netwerk.

c. (4 punten) Neem aan dat F een *Noisy-OR*-knoop is, waarbij geldt: $P(f|s, \neg w, \neg a) = c_1$, $P(f|\neg s, w, \neg a) = c_2$ en $P(f|\neg s, \neg w, a) = c_3$. We gebruiken hier de vertrouwde notatie: kleine letters staan voor het waar zijn van de variabele met de bijbehorende hoofdletter. Wat zijn de andere relevante kansen $P(f|...)$, uitgedrukt in c_1 , c_2 en c_3 ?

d. (4 punten) Geef de vier soorten queries, en van elk een voorbeeld uit het netwerk. Eén van de voorbeelden moet de query van **e** zijn.

e. (5 punten) Druk de kans dat het slecht weer is, gegeven dat er een file staat maar dat het geen spitsuur is, uit in “bekende” kansen.

Opgave 5: Theorie (diversen) (20 punten)

a. (8 punten) Geef een *P/E/A/S-omschrijving* van een computersysteem dat controleert of er bij een voetbalwedstrijd een doelpunt is gemaakt. Vermeld daartoe waar de vier letters voor staan, en geef per letter enkele steekwoorden.

b. (6 punten) Wat zijn trainingsset, validatieset en testset bij een lerend algoritme, en waarom zijn ze (in principe) nodig?

c. (6 punten) We willen een probleem oplossen met een *Genetisch Algoritme*. In het algemeen kan bij mutatie van een individu wellicht een individu ontstaan dat geen (serieuze) kandidaat-oplossing representeert. Geef drie manieren om deze moeilijkheid te benaderen.