

Tentamen IN2205 Kennissystemen

19 Januari 2009, 14:00–17:00

- Dit tentamen heeft 6 meerkeuzevragen in totaal goed voor 12 punten en 7 open vragen met in totaal 39 punten.
- Wat betreft de meerkeuzevragen:
 - Er is voor iedere vraag telkens maar één goed antwoord mogelijk.
- Wat betreft de open vragen:
 - Geef antwoord in correct Nederlands of Engels en schrijf leesbaar (gebruik eerst kladpapier).
 - Motiveer je antwoorden.
 - Geef geen irrelevante informatie. Dit kan leiden tot puntenaftrek.

Vraag:	7	8	9	10	11	12	13	Totaal:
Punten:	10	3	4	6	9	4	3	39

- Het gebruik van boek of aantekeningen tijdens dit tentamen is niet toegestaan.
- Het gebruik van een rekenmachine is toegestaan.
- Voordat je je antwoorden inlevert, controleer of op ieder blaadje je naam en studienummer staat en geef het aantal ingeleverde bladen aan op (tenminste) de eerste pagina.
- De tentamenstof bestaat uit hoofdstukken 1,2,3,4,6,7,8 en 9 uit *Artificial Intelligence* van Michael Negnevitsky, behalve secties 3.4, 3.5, 8.2, 9.1 en 9.7, het artikel *Swarm Smarts* van Eric Bonabeau *et. al.*, de lecture notes over Bayesian Reasoning en Swarm Intelligence en de handouts over reinforcement learning.
- Uiteraard komen in één tentamen niet alle onderwerpen aan bod. Trek daarom op basis van dit tentamen geen conclusies over stof die nooit getoetst wordt.
- Totaal aantal pagina's: 7.

Succes!

Multiple-choicevragen

1. (2 punten) Een certainty factor is . . .
- A. de kans van een regel in een expertsysteem.
 - B. het vertrouwen dat een expert stelt in een regel in een expertsysteem.**
 - C. de prioriteit van een regel in een expertsysteem.
 - D. de mate van vaagheid van een regel in een expertsysteem.

2. (2 punten) Stel we hebben de fuzzy sets A , B met membershipfuncties:

$$\mu_A(x) = (0.1/1, 0.2/2, 0.9/3, 0.7/4, 0.3/5) \quad (1)$$

$$\mu_B(x) = (0.7/1, 0.4/2, 0.8/3, 0.2/4, 0.6/5) \quad (2)$$

Wat is de membershipfunctie van de fuzzy set $\neg(\neg A \cup B)$ als we gebruik maken van de standaardoperatoren?

- A. $(0.7/1, 0.4/2, 0.1/3, 0.2/4, 0.6/5)$
 - B. $(0.1/1, 0.2/2, 0.2/3, 0.7/4, 0.3/5)$**
 - C. $(0.9/1, 0.8/2, 0.8/3, 0.3/4, 0.7/5)$
 - D. $(0.3/1, 0.6/2, 0.9/3, 0.8/4, 0.4/5)$
3. (2 punten) De leerregel van Hebb voor neurale netwerken luidt:

$$\Delta w_{ij}(p) = \alpha y_j(p) x_i(p) - \phi y_j(p) w_{ij}(p). \quad (3)$$

Waarvoor staat ϕ in deze formule?

- A. de learning rate
 - B. het momentum
 - C. een inertia weight
 - D. de forgetting factor**
4. (2 punten) Een self-organising map (Kohonen netwerk) maakt gebruik van . . .
- A. supervised learning.
 - B. reinforcement learning.
 - C. unsupervised learning.**
 - D. geen van bovenstaande antwoorden.
5. (2 punten) Wat is een verschil tussen evolution strategies en genetic algorithms? Evolution strategies gebruiken geen . . .
- A. fitnessfunctie.
 - B. codering van de parameters.**
 - C. selectiemechanisme.
 - D. mutatie.

6. (2 punten) Bij particle swarm optimization (PSO) wordt de snelheid van een particle aangepast volgens:

$$v_{ij}(t+1) = \phi v_{ij}(t) + c_1 r_{1j} [y_{ij}(t) - x_{ij}(t)] + c_2 r_{2j} [\hat{y}_{ij}(t) - x_{ij}(t)]. \quad (4)$$

Stel dat de beginposities van de particles gelijkmatig verdeeld zijn over de oplossingsruimte. Om te zorgen dat het algoritme niet te snel convergeert naar een lokaal optimum moeten we zorgen dat ...

- A. c_1 niet te groot is t.o.v. andere parameters.
- B. c_2 niet te groot is t.o.v. andere parameters.**
- C. ϕ niet te groot is t.o.v. andere parameters.
- D. $y_{ij}(t)$ niet te groot is t.o.v. andere parameters.

Open vragen

7. Geef voor ieder van de onderstaande vraagstukken aan welke van de in dit vak behandelde oplossingsmethoden je zou gebruiken, motiveer je antwoorden.

- (a) (2 punten) Spellingscorrectie.

Antwoord: Een Hopfield neuraal netwerk kan fouten corrigeren en heeft als voordeel dat tijdens de training alleen voorbeelden van correcte woorden nodig zijn.

Een expertsysteem zou ook kunnen. Met name voor spelfoutdetectie (hoewel opzoeken in een woordenboek dan handiger is) maar heeft als nadeel dat het lastig eenduidig te bepalen is welk correct woord bij een fout hoort. In praktijk worden hiervoor kansmodellen gebruikt die naar de omringende woorden kijken om te bepalen wat het meest waarschijnlijke correcte woord is.

- (b) (2 punten) Uit een verzameling van munten het minimale aantal munten selecteren waarmee je een gegeven bedrag gepast kunt betalen (je mag aannemen dat er altijd een oplossing is).

Antwoord: Dit is een optimaliseringsprobleem, een variant van het zogenaamde knapsackprobleem dat vergelijkbaar is met het traveling salesmanprobleem. Particle swarm optimization, ant-based optimization en evolutionary computation liggen dus voor de hand. Let op: de simpele strategie om steeds de munt met de grootst mogelijke waarde te nemen gaat niet altijd op. Stel bijvoorbeeld dat je 2.15 moet betalen terwijl je 1 2 euro munt hebt, 1 euro munt, een munt van 50 cent, 3 munten van 20 cent en een munt van 5 cent, dan kun je de 2 euro munt niet gebruiken.

- (c) (2 punten) een context-gevoelige hulpfunctie voor een tekstverwerker, die aan hand van wat je aan het doen bent zo relevant mogelijke suggesties doet zodra je om hulp vraagt.

Antwoord: Hoewel expertsystemen geschikt zijn voor het implementeren van hulpfuncties, hebben we hier het bijkomstige probleem dat het systeem contextgevoelig moet zijn, een systeem dat overweg kan met ontbrekende, afwijkende of onzekere informatie en dat het verband tussen situaties en de vereiste hulp kan leren, zoals een neuraal netwerk of een Bayesiaans netwerk verdient dus de voorkeur.

- (d) (2 punten) Het plannen van een kerstdiner voor de bewoners van een studentenhuus, zodanig dat zoveel mogelijk bewoners 's avonds mee kunnen eten en dat iedereen een gang kan klaarmaken. Sommige gangen moeten op de avond zelf klaargemaakt worden en andere kunnen eerder voorbereid worden. De gangen moeten zo over de bewoners verdeeld worden dat dit binnen hun agenda past (denk aan colleges, tentamenperiodes en uitslapen na een zware nacht).

Antwoord: Dit komt neer op het maken van een rooster met een flink aantal eisen en beperkingen, genetische algoritmen zijn daar uitermate geschikt voor.

- (e) (2 punten) Een intelligente cruise control die gasgeven, remmen en schakelen in een auto voor zijn rekening neemt met behulp van informatie over de door de bestuurder gewenste snelheid, de toegestane snelheid, de afstand en snelheid t.o.v. omringende auto's, de huidige versnelling, het aantal toeren dat de motor maakt enzv.

Antwoord: De benodigde kennis valt uit te drukken in regels. Aangezien we te maken hebben met continue waarden afkomstig van (mogelijk onnauwkeurige) sensoren is een fuzzy expertsysteem dan de beste keuze. Als alternatief zou je voor een neurale netwerk kunnen kiezen dat dan door een bestuurder getraind moet worden.

8. (3 punten) Stel, we hebben een knowledge base met de volgende regels:

1. IF a THEN b
2. IF t AND s THEN z
3. IF w THEN v
4. IF c AND r THEN q
5. IF b AND q AND h THEN p
6. IF t THEN a
7. IF s THEN t
8. IF h AND z THEN c
9. IF f AND j THEN d
10. IF y and u THEN w
11. IF g THEN h
12. IF z THEN q

In de database komen de volgende feiten voor: **s, r, g**.

We gaan uit van een seriële (depth-first) verwerking van regels. Laat zien hoe de inference engine alle feiten afleidt die kunnen worden afgeleid uit de database. Geef per *cycle* aan welke regels vuren (in de juiste volgorde) en hoe de database verandert.

Antwoord:

Regel vuurt	database
7	g, r, s, t
11	g, h , r, s, t
2	g, h, r, s, t, z
6	a , g, h, r, s, t, z
8	a, c , g, h, r, s, t, z
12	a, c, g, h, q , r, s, t, z
1	a, b , c, g, h, q, r, s, t, z
4	a, b, c, g, h, q, r, s, t, z
5	a, b, c, g, h, p , q, r, s, t, z

Meer regels kunnen niet vuren.

9. (a) (2 punten) Leg uit hoe defuzzificatie in een fuzzy expertsysteem werkt en waarom dit nodig is. Ga er in je antwoord vanuit dat het systeem gebruik maakt van Mamdani inferentie.

Antwoord: Defuzzificatie zet de uitkomst van het fuzzy expertsysteem om van een fuzzy set naar een waarde. Dit kan met de Center of Gravity methode (zie boek, blz. 111). De eigenlijke waarde wordt geschat door de lidmaatschapswaarde op een aantal punten te meten en deze te gebruiken om een gewogen gemiddelde te berekenen.

- (b) (2 punten) Wat is het verschil tussen de membershipwaarde in een fuzzy set en een kans?

Antwoord: Een membershipwaarde geeft aan in hoeverre een object of gebeurtenis deel uitmaakt van een verzameling. Voor een regenbui kun je bijvoorbeeld afvragen in hoeverre hier sprake is van motregen. Een kans geeft aan hoe waarschijnlijk het is dat iets gebeurt, bijvoorbeeld dat het straks gaat regenen. Hier is dus geen onzekerheid over de aard van de gebeurtenis (het regent of het regent niet).

10. (a) (3 punten) Beschrijf de structuur en de werking van een Hopfield neural network (je hoeft geen formules te geven, dit mag wel).

Antwoord: Een Hopfield netwerk is een recurrent netwerk. De outputs worden weer teruggevoerd naar de inputs. Preciezer, de outputs van alle andere neuronen dienen als input van een neuron (geef dit eventueel weer in een plaatje). Dit gaat net zo lang door totdat het netwerk een stabiele toestand bereikt, input en output zijn dan aan elkaar gelijk (maar niet per se aan de oorspronkelijke input).

Deze stabiele toestanden worden wel fundamental memories genoemd. We kunnen het netwerk een aantal van zulke memories laten leren. Hier is een maximum aan (de formule van het maximum aantal memories geeft extra punten). Het netwerk beeld een input af op een fundamental memory, dit is echter niet altijd het dichtstbijzijnde element in het geheugen.

De activatiefunctie is de sign-functie, met als bijzondere eigenschap dat de state gelijk blijft als de gewogen input van het neuron nul is.

De leerregel is een variant van de leerregel van Hebb, het gewicht wordt groter als de twee neuronen verbonden met dat gewicht vaak tegelijk dezelfde waarde hebben (of juist een tegengestelde waarde, dan wordt het gewicht negatief).

- (b) (1 punt) Noem een mogelijke toepassing van Hopfield netwerken, motiveer je antwoord.

Antwoord: Foutcorrectie, patroonherkennen (handschriften, spraak), classificatie.

- (c) (2 punten) Wat is het verschil tussen supervised learning en reinforcement learning?

Antwoord: Bij supervised learning worden voorbeelden van de juiste output gebruikt, bij reinforcement learning wordt slechts een waardering voor de output/acties gegeven.

11. In een dierentuin wil men een fuzzy expertsysteem ontwikkelen dat de omstandigheden in een groot aquarium optimaal houdt. Als eerste stap heeft men nauwkeurig bijgehouden hoe de medewerkers van het aquarium dit doen. Dit heeft geresulteerd in een data set van metingen met daarbij de juiste maatregelen. Hieruit zijn een aantal fuzzy regels afgeleid. Bijvoorbeeld:

- IF *CO2-niveau* is laag AND *zuurstof-niveau* is laag THEN *CO2-toevoer* is groot
- IF *CO2-niveau* is laag AND *zuurstof-niveau* is normaal THEN *CO2-toevoer* is beperkt
- IF *nitraat-niveau* is hoog THEN *water-verversing* is veel

- (a) (8 punten) Als laatste stap moeten de parameters van de fuzzy sets, *laag*, *hoog*, *groot* enzovoorts, bepaald worden. Hoe zouden we dit kunnen doen met behulp van genetische algoritmen? Behandel in je antwoord:

- de representatie als chromosoom,

Antwoord: Een fuzzy set kan worden weergegeven door een functie, als we uitgaan van standaard 'driehoekige' fuzzy set dan zijn de parameters van de functie de hoogte en de breedte (zie boek blz. 271) als alternatief kunnen we simpelweg het middelpunt en de afstand tot het beginpunt van de driehoek geven of een lijst van hoekpunten. In al deze gevallen wordt een fuzzy set dus door een reeks getallen weergegeven. Deze getallen kunnen we als bitstrings coderen die samen het chromosoom vormen. Het is ook mogelijk de getallen zelf als de genen in het chromosoom te zien, dit vereist wat extra aandacht bij cross-over en mutatie.

- de populatie,

Antwoord: N willekeurig gekozen chromosomen.

- fitness,

Antwoord: Ieder chromosoom representeert de parameters van het fuzzy expertsysteem. Aangezien een data set van metingen (inputs) met juiste maatregelen (outputs) verzameld is, kunnen we per chromosoom testen hoe goed het fuzzy expertsysteem werkt met deze parameters door de inputs in te voeren en de verkregen uitkomsten te vergelijken met de gewenste uitkomsten, bijvoorbeeld met een mean-squared error.

- selectie,

Antwoord: We selecteren chromosomen voor cross-over en mutatie met een kans proportioneel met hun fitness. Een methode hiervoor is 'roulette wheel selection'. De oppervlakte behorende bij een chromosoom op het roulettewiel komt overeen met de fitness ratio, dit is de fitness van het chromosoom gedeeld door de som van de fitnesswaarden van alle chromosomen.

- cross-over,

Antwoord: Bij cross-over worden twee chromosomen geselecteerd. Er wordt een breekpunt gekozen tussen twee willekeurige genen de genen (bits) achter dit breekpunt van het ene chromosoom worden verwisseld met de genen achter het breekpunt van het andere chromosoom (je kunt dit het handigst met een plaatje weergeven). Cross-over vindt plaats met een bepaalde kans, deze heeft meestal een waarde van 0.7. Als geen cross-over plaatsvindt worden de ouder chromosomen simpelweg doorgegeven (cloning).

- mutatie,

Antwoord: Met een kleine kans (tussen de 0.01 en de 0.001) wordt een bit in de nieuwe chromosomen verwisseld (0 wordt 1 en andersom). Als je er voor gekozen hebt de parameters zelf als genen te zien dan is mutatie mogelijk door een kleine waarde bij de parameter op te tellen, deze waarde wordt dan bepaald aan hand van een kansverdeling.

- de nieuwe populatie en terminatie.

Antwoord: Selectie, cross-over en mutatie worden doorlopen totdat de omvang van de nieuwe populatie N is. De nieuwe populatie vervangt dan de oude populatie in de volgende iteratie. (Dit gaat een van te voren bepaald aantal iteraties door of totdat weinig verandering meer optreedt.)

- (b) (1 punt) Hoe zouden we het genetische algoritme kunnen gebruiken om niet alleen de parameters van de fuzzy sets te leren, maar ook de fuzzy regels zelf?

Antwoord: We kunnen aan de chromosomen bits toevoegen die aangeven of een regel wel of niet gebruikt moeten worden in het fuzzy expertsysteem. Je zou hier ook voor een genetic programming aanpak kunnen kiezen, waarbij de fuzzy regels in de vorm van een boom worden gecodeerd met bijvoorbeeld eerst een laag if knopen die een antecedent en een consequent hebben, de antecedent kan dan operator knopen als kinderen hebben (and, or, not) die op hun beurt fuzzy set knopen als kinderen hebben.

12. (4 punten) Het vriest al een aantal dagen aan een stuk en de heer Holmes heeft gehoord dat schaatsen op 70% van alle slootjes mogelijk is. Hij vraagt zich nu af of hij kan schaatsen op de sloot die achter zijn huis langs loopt. Hij besluit om te kijken of hij op het ijs bij de stijger net achter zijn huis kan staan. Als de hele sloot voldoende bevroren is dan moet dit zeker lukken. Als de sloot echter niet overal voldoende bevroren is dan gaat dit 60% van de tijd toch nog goed. Dan ziet Holmes het zoontje van buurman Watson op het ijs. Als de sloot zo goed bevroren is dat Holmes kan erop kan schaatsen, dan kan het ijs het kleine jongetje zeker houden, maar als de sloot niet zodanig bevroren is dat Holmes kan gaan schaatsen dan is er toch nog een kans van 50% dat het lichtere buurjongetje er wel op kan schaatsen. Geven dat Holmes bij de stijger achter zijn huis op het ijs kan staan en dat het buurjongetje vrolijk rondjes schaatst over de sloot, wat is de kans dat Holmes op de sloot kan schaatsen?

Antwoord: Uit de tekst kunnen we de volgende kansen halen:

$$P(S = true) = 0.7, \quad (5)$$

waarbij S staat voor 'Holmes kan op de sloot schaatsen',

$$P(S = true|H = true) = 1.0, \quad (6)$$

$$P(S = true|H = false) = 0.6, \quad (7)$$

waarbij H staat voor 'Holmes kan staan op het ijs bij de steiger achter zijn huis' en

$$P(B = true|S = true) = 1.0, \quad (8)$$

$$P(B = true|S = false) = 0.5, \quad (9)$$

waarbij B staat voor 'het ijs is stevig genoeg voor het buurjongetje om op te schaatsen'.

Gevraagd is de kans dat Holmes kan gaan schaatsen:

$$(10)$$

$P(S = true|H = true, B = true)$.

We kunnen dit berekenen door de regel van Bayes toe te passen (en conditionele onafhankelijkheid te veronderstellen):

$$P(S = true|H = true, B = true) = \alpha P(H = true|S = true)P(B = true|S = true)P(S = true) \quad (11)$$

met

$$\alpha = P(H = true|S = true)P(B = true|S = true)P(S = true) + P(H = true|S = false)P(B = true|S = false)P(S = false) \quad (12)$$

Invullen van de waarden geeft:

$$\frac{1.0 * 1.0 * 0.7}{1.0 * 1.0 * 0.7 + 0.6 * 0.5 * 0.3} = 0.886 \quad (13)$$

13. (3 punten) Stel we hebben een simpel wegennetwerk waarin een stad A verbonden is met steden B en C die op hun beurt allebei verbonden zijn met stad D. Leg uit hoe ant-based algorithms in staat zijn om de kortste route van A naar D te bepalen.

Antwoord: In het begin volgt een mier een willekeurige route (dus A-B-D of A-C-D). De hoeveelheid feromoon die aan de verbindingen langs deze route wordt toegekend is omgekeerd evenredig met de afgelegde afstand. Volgende mieren kiezen hun route met een kans die evenredig is met de hoeveelheid feromoon. De kortere route heeft dus een grotere kans om gekozen te worden. Doordat de kortere route vaker gekozen wordt, neemt de hoeveelheid feromoon langs deze route veel sneller toe dan voor de langere route, waardoor de kans dat deze route gekozen wordt nog verder toeneemt. Op den duur volgen vrijwel alle mieren de kortere route.

Einde tentamenopgaven.

Controleer voor de zekerheid of je alle vragen hebt beantwoord. Het zouden er 13 moeten zijn.