

Tentamen IN2205 Kennissystemen

21 Januari 2008, 14:00–17:00

- Het gebruik van boek of aantekeningen tijdens dit tentamen is niet toegestaan.
- Het gebruik van een rekenmachine is toegestaan.
- Dit tentamen heeft 9 meerkeuzevragen in totaal goed voor 10 punten en 6 open vragen met in totaal 40 punten. Je cijfer voor dit tentamen wordt berekend door het aantal behaalde punten door 5.0 te delen. Het cijfer voor dit tentamen telt voor $\frac{8}{10}$ mee in je eindcijfer. De rest van je cijfer wordt bepaald uit de cijfers voor de opdrachten.
- 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.
- 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).
 - Geef geen irrelevante informatie. Dit kan leiden tot puntenaftrek.

Vraag:	10	11	12	13	14	15	Totaal:
Punten:	4	6	8	8	9	5	40

- 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 Chapters 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. en de lecture notes over Bayesian Reasoning en Swarm Intelligence.
- Totaal aantal pagina's: 8.

Succes!

1. (1 punt) Welke uitspraak omtrent regelgebaseerde expert systemen is juist?
- A. Expert systemen gebruiken conflict resolution om met onvolledige en vage data overweg te kunnen.
 - B. Expert systemen zijn zelf-organiserende systemen.
 - C. Expert systemen leren automatisch nieuwe regels van voorbeelden met behulp van backward chaining.
 - D. Expert systemen kunnen een verklaring geven voor de gevonden oplossing.

Antwoord: D. Expert systemen kunnen een verklaring geven voor de gevonden oplossing.

2. (1 punt) Waarvoor gebruiken expert systemen conflict resolution?
- A. Om te bepalen welke regel mag vuren als meer dan één regel kan vuren.
 - B. Om te bepalen of forward chaining of backward chaining gebruikt moet worden.
 - C. Om te bepalen welke consequent de grootste kans heeft.
 - D. Om te bepalen welke regels tot een bepaalde conclusie hebben geleid.

Antwoord: A.

Om te bepalen welke regel mag vuren als meer dan een regel kan vuren.

3. (1 punt) Wat is een manier om de vereniging (OR) van twee fuzzy sets te definiëren?
- A. $\mu_{A \cup B}(x) = \min[\mu_A(x), \mu_B(x)]$.
 - B. $\mu_{A \cup B}(x) = \mu_A(x) \times \mu_B(x)$.
 - C. $\mu_{A \cup B}(x) = \mu_A(x) + \mu_B(x) - \mu_A(x) \times \mu_B(x)$.
 - D. Geen van bovenstaande antwoorden.

Antwoord: C.

$\mu_{A \cup B}(x) = \mu_A(x) + \mu_B(x) - \mu_A(x) \times \mu_B(x)$.

4. (2 punten) Stel we hebben de fuzzy sets A , B en C met membershipfuncties:

$$\mu_A(x) = (0.8/1, 0.2/2, 0.7/3, 0.1/4, 0.9/5, 0.2/6, 1.0/7, 0.0/8, 1.0/9) \quad (1)$$

$$\mu_B(x) = (0.1/1, 0.2/2, 0.3/3, 0.4/4, 0.5/5, 0.6/6, 0.7/7, 0.8/8, 1.0/9) \quad (2)$$

$$\mu_C(x) = (1.0/1, 0.9/2, 0.8/3, 0.7/4, 0.6/5, 0.5/6, 0.4/7, 0.3/8, 0.2/9) \quad (3)$$

Wat is de membershipfunctie van de fuzzy set A of *niet* (B en C) als we gebruik maken van de standaard-operatoren?

- A. $\mu(x) = (0.9/1, 0.8/2, 0.7/3, 0.6/4, 0.5/5, 0.5/6, 0.6/7, 0.7/8, 0.8/9)$.
- B. $\mu(x) = (0.0/1, 0.1/2, 0.2/3, 0.1/4, 0.4/5, 0.2/6, 0.3/7, 0.0/8, 0.0/9)$.
- C. $\mu(x) = (1.0/1, 0.9/2, 0.8/3, 0.9/4, 0.6/5, 0.8/6, 0.7/7, 1.0/8, 1.0/9)$.
- D. $\mu(x) = (0.9/1, 0.8/2, 0.7/3, 0.6/4, 0.9/5, 0.5/6, 1.0/7, 0.7/8, 1.0/9)$.

Antwoord: D.

volgt uit $\max(\mu_a, 1 - \min(\mu_b, \mu_c))$

5. (1 punt) Sugeno fuzzy inferentie verschilt van Mamdani fuzzy inferentie ...
- A. omdat Sugeno geen defuzzificatie-stap nodig heeft en Mamdani wel.
 - B. omdat Mamdani geen defuzzificatie-stap nodig heeft en Sugeno wel.
 - C. omdat de consequent van een regel bij Sugeno een functie is van de inputs en bij Mamdani een fuzzy set.
 - D. omdat de consequent van een regel bij Mamdani een functie is van de inputs en bij Sugeno een fuzzy set.

Antwoord: C.

omdat de consequent van een regel bij Sugeno een functie is van de inputs en bij Mamdani een fuzzy set.

6. (1 punt) Wat is een verschil tussen kansrekening en fuzzy logic?
- A. Kansrekening kwantificeert nauwkeurigheid, fuzzy logic kwantificeert zekerheid.
 - B. Kansrekening kwantificeert zekerheid, fuzzy logic kwantificeert nauwkeurigheid.
 - C. Kansrekening kwantificeert onzekerheid, fuzzy logic kwantificeert zekerheid.
 - D. Kansrekening kwantificeert onnauwkeurigheid, fuzzy logic kwantificeert nauwkeurigheid.

Antwoord: B.

Kansrekening kwantificeert zekerheid, fuzzy logic kwantificeert nauwkeurigheid.

7. (1 punt) Waarvoor staat β in de *generalised delta rule* voor het trainen van neurale netwerken:

$$\Delta w_{jk}(p) = \beta \times \Delta w_{jk}(p-1) + \alpha \times y_j(p) \times \delta_k(p)? \quad (4)$$

- A. De learning rate, een positieve constante kleiner dan 1.
- B. De threshold, een constante tussen -1 en 1.
- C. Het momentum, een positieve constante kleiner dan 1.
- D. De gradient, een constante tussen -1 en 1.

Antwoord: C.

Het momentum, een positieve constante kleiner dan 1.

8. (1 punt) Het Hopfield neurale netwerk heeft als eigenschap dat ...
- A. voor training voor iedere input de gewenste output bekend moet zijn.
 - B. het aantal fundamentele geheugens beperkt is.
 - C. de Euclidische afstand als activatiefunctie gebruikt wordt.
 - D. fouten van achter naar voor door het netwerk propageren om de gewichten te trainen.

Antwoord: B.

het aantal fundamentele geheugens beperkt is.

9. (1 punt) Bij ant-based optimization ...

- A. probeert iedere mier een optimaal pad naar de bestemming te vinden.
- B. wordt met behulp van de fitnessfunctie bepaald welke mieren doorgaan in de volgende generatie.
- C. is de eindoplossing niet terug te voeren op het gedrag van een individuele mier.
- D. wordt na verloop van tijd altijd het optimale pad gevonden.

Antwoord: C.

is de eindoplossing niet terug te voeren op het gedrag van een individuele mier.

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

(a) (1 punt) Advies bij het invullen van een belastingformulier.

Antwoord: Expert systeem, want alle informatie is in principe beschikbaar in wet- en regelgeving, alleen erg ingewikkeld.

(b) (1 punt) In een autofabriek zijn een aantal assemblagerobots beschikbaar. We willen nu weten in welke volgorde deze robots langs de assemblage lijnen opgesteld moeten worden om zoveel mogelijk auto's in zo kort mogelijke tijd te produceren. Van sommige robots zijn meerdere exemplaren beschikbaar zodat sommige stukken van de assemblage in parallel kunnen worden uitgevoerd. Van tevoren is bekend welke stappen van andere stappen afhangen (je hebt bijvoorbeeld een carrosserie nodig om er deuren aan te kunnen hangen) en hoeveel tijd een robot nodig heeft om een taak uit te voeren.

Antwoord: Dit is een complex optimaliseringsprobleem. Genetische algoritmen of eventueel particle swarm optimization.

(c) (1 punt) Een spamfilter.

Antwoord: Een neuraal netwerk of Bayesiaans systeem dat aan hand van voorbeelden van spam en e-mails leert wat spam is. Inputs zouden bijvoorbeeld woorden of woordcombinaties in het subject of in de mail zelf kunnen zijn.

(d) (1 punt) Een systeem dat hulpverleners bij humanitaire rampen helpt bij het kiezen van type noodwoningen (bijvoorbeeld tenten, houten barakken, containers, ...) aan hand van gegevens over klimaat, seizoen, geografische ligging, aantal vluchtelingen en termijn waarop de woningen af moeten zijn enzovoorts.

Antwoord: Gezien de aard van de inputs ligt een fuzzy systeem hier voor de hand. Je zou echter ook aan een neuraal netwerk kunnen denken dat getraind wordt op voorbeelden van eerdere rampen.

11. Stel, we hebben een knowledge base met de volgende regels:

1. IF a AND d THEN e
2. IF a THEN k
3. IF u AND c THEN t
4. IF i THEN h
5. IF b THEN u
6. IF r THEN t
7. IF c THEN f
8. IF c THEN q
9. IF t THEN d
10. IF i THEN g
11. IF k THEN p
12. IF g OR f THEN r
13. IF e AND t THEN s
14. IF p AND q THEN x
15. IF x AND b THEN y
16. IF r AND s THEN z

In de database komen de volgende feiten voor: **a, b, c**.

We gaan uit van een seriële (depth-first) verwerking van regels.

- (a) (3 punten) Laat zien hoe de inference engine aantoonst of doel z waar is. Geef aan welke feiten op de stack worden geplaatst en hoe de database verandert.

Antwoord: Database: a, b, c.

Doelen: z.

Stack: 16. Doelen r, s, z.

Stack: 12, 16. Doelen: g (f), r, s, z.

Stack: 10, 12, 16. Doelen: i, g (f), r, s, z.

Er is geen regel die i kan vervullen en geen andere manier om g te krijgen. Stack: 12, 16. Doelen: f, r, s, z.

Regel 7 kan vuren, voegt f toe aan database.

Hierdoor kan ook regel 12 vuren, voegt r toe aan database.

Stack: 16. Doelen: s, z.

Stack: 13, 16. Doelen: e, t, s, z.

Stack: 1, 13, 16. Doelen: d, e, s, z.

Stack: 9, 1, 13, 16. Doelen: t, d, e, s, z.

Stack: 3, 9, 1, 13, 16. Doelen: u, t, d, e, s, z.

Regel 5 vuurt, want b in database, voegt u toe aan database.

Regel 3 vuurt, voegt t toe aan database.

Stack: 9, 1, 13, 16. Doelen: d, e, s, z.

Regel 9 vuurt, voegt d toe aan database.

Stack: 1, 13, 16. Doelen: e, s, z.

Regel 1 vuurt, voegt e toe aan database.

Stack: 13, 16. Doelen: s, z.

Regel 13 vuurt, voegt s toe aan database.

Stack: 16. Doelen: z.

Regel 16 vuurt, voegt z toe aan database.

Doel vervuld.

- (b) (3 punten) Laat zien hoe de inference engine alle feiten afleidt die kunnen worden afgeleid. Geef per cycle aan welke regels vuren en hoe de database verandert.

Antwoord:

Regel vuurt	database
2	a, b, c, k
5	a, b, c, k, u
7	a, b, c, f , k, u
8	a, b, c, f, k, q , u
11	a, b, c, f, k, p , q, u
12	a, b, c, f, k, p, q, r , u
14	a, b, c, f, k, p, q, r, u, x
15	a, b, c, f, k, p, q, r, u, x, y
3	a, b, c, f, k, p, q, r, t, u, x, y
6	a, b, c, f, k, p, q, r, t, u, x, y
9	a, b, c, d , f, k, p, q, r, t, u, x, y
1	a, b, c, d, e , f, k, p, q, r, t, u, x, y
13	a, b, c, d, e, f, k, p, q, r, s , t, u, x, y
16	a, b, c, d, e, f, k, p, q, r, s, t, u, x, y, z

Meer regels kunnen niet vuren.

12. In deze opgave gaan we de kans berekenen dat iemand dit tentamen haalt. We geven 'halen van het tentamen' weer met de stochastische variabele T . Zonder enige voorbereiding is de kans op het halen van het tentamen 0.05. Uit voorgaande jaren is echter gebleken dat mensen die het tentamen halen met 70% kans alle colleges gevolgd hebben. Mensen die het tentamen niet halen hebben juist met 95% kans niet alle colleges gevolgd. Degene die het tentamen halen hebben met 90% kans het boek bestudeerd en met 80% kans oefententamen gemaakt. Degene die het tentamen niet halen hebben met 50% kans het boek bestudeerd en met 75% kans oefenopgaven gemaakt.

De stochastische variabele C geeft aan of iemand colleges heeft gevolgd, B of iemand het boek heeft bestudeerd en O of hij/zij de oefenopgaven heeft gemaakt. We veronderstellen (misschien niet geheel terecht) dat C , B en O onafhankelijk zijn gegeven de informatie of iemand het tentamen haalt of niet (T).

- (a) (4 punten) Geef de kanstabellen voor $P(T)$, $P(C|T)$, $P(B|T)$, $P(O|T)$.

Antwoord:

$$P(T) = \begin{array}{|c|c|} \hline T & \\ \hline false & 0.95 \\ true & 0.05 \\ \hline \end{array} \quad P(C|T) = \begin{array}{|c|c|c|} \hline C & T & \\ \hline false & false & 0.95 \\ true & false & 0.05 \\ false & true & 0.30 \\ true & true & 0.70 \\ \hline \end{array} \quad P(B|T) = \begin{array}{|c|c|c|} \hline B & T & \\ \hline false & false & 0.5 \\ true & false & 0.5 \\ false & true & 0.1 \\ true & true & 0.9 \\ \hline \end{array} \quad (5)$$

$$P(O|T) = \begin{array}{|c|c|c|} \hline O & T & \\ \hline false & false & 0.25 \\ true & false & 0.75 \\ false & true & 0.2 \\ true & true & 0.8 \\ \hline \end{array} \quad (6)$$

- (b) (4 punten) Wat is de kans dat iemand die niet alle colleges gevolgd heeft en het boek niet heeft bestudeerd maar wel opgaven heeft gemaakt het tentamen haalt? Laat je berekening zien.

Antwoord: We zoeken $P(T = true|C = false, B = false, O = true)$, toepassen van de regel van Bayes' en onafhankelijkheid tussen de 'symptomen' geeft:

$$P(T|C, B, O) = \frac{P(C|T)P(B|T)P(O|T)P(T)}{\sum_T P(C|T)P(B|T)P(O|T)P(T)} = \alpha P(C|T)P(B|T)P(O|T)P(T) \quad (7)$$

$$\alpha P(T|C = false, B = false, O = true) = \begin{array}{|c|c|} \hline T & \\ \hline false & 0.95 \times 0.5 \times 0.75 \times 0.95 = 0.3384 \\ true & 0.3 \times 0.1 \times 0.8 \times 0.05 = 0.0012 \\ \hline \end{array} \quad (8)$$

$$\alpha = 1/(0.3384 + 0.0012) = 1/0.3396 \quad (9)$$

$$P(T|C = false, B = true, O = false) = \begin{array}{|c|c|} \hline T & \\ \hline false & 0.3384/0.3396 = 0.9964 \\ true & 0.0012/0.3396 = 0.0035 \\ \hline \end{array} \quad (10)$$

De gezochte kans is dus 0.0035.

13. Het perceptron is het meest eenvoudige neurale netwerk. De leerregel voor het perceptron luidt:

$$w_i(p+1) = w_i + \alpha \times x_i(p) \times e(p). \quad (11)$$

- (a) (4 punten) Beschrijf het perceptron. Behandel (minimaal) de netwerkstructuur, de inputs, de outputs en de activatiefunctie.

Antwoord: 1 neuron, n inputs, 1 output, activatiefunctie is een signfunctie (geeft +1 of -1).

- (b) (3 punten) Leg uit hoe leren in een perceptron werkt, geef daarbij aan waar de termen in bovenstaande leerregel voor staan.

Antwoord: Aanpassen van de gewichten aan hand van de gemaakte fout $e(p)$ die het verschil geeft tussen de gewenste en de werkelijke output. w_i is de vorige waarde van het gewicht behorende bij input x_i . De verandering van het gewicht is proportioneel met de fout en met de input waarde. α is de learning rate, een getal tussen 0 en 1 dat bepaald hoe snel gewichten worden aangepast.

(c) (1 punt) Welke functies kan een perceptron wel/niet leren?

Antwoord: Perceptron legt een lineaire scheiding aan.

14. Genetische algoritmen kunnen gebruikt worden voor het trainen van neurale netwerken.

(a) (7 punten) Leg uit hoe dit in zijn werk gaat. Ga uit van een multilayer feedforward netwerk met een of meer hidden layers. Behandel daarbij:

- de representatie als chromosoom,
- de populatie,
- fitness,
- selectie,
- cross-over,
- mutatie,
- de nieuwe populatie en terminatie.

Antwoord: Het chromosoom is een vector van gewichten (als we de gewichten willen updaten) of van 1-en en nullen als we alleen de structuur willen updaten.

De populatie bestaat uit een aantal random gegenereerde chromosomen.

De fitness kan gemeten worden door voor voor ieder chromosoom in de populatie het bijbehorende netwerk te testen op een verzameling testvoorbeelden waarvan de gewenste output bekend is. De mean squared error op deze set is dan de fitness (hoe dichterbij nul hoe beter).

Selectie kan geschieden met behulp van roulette wheel selection. De kans dat een chromosoom geselecteerd wordt voor cross-over is proportioneel met de fitness. Hiervoor berekenen we de fitness-ratio van een chromosoom: zijn fitness gedeeld door de som van alle fitness waarden.

bij cross-over worden twee chromosomen geselecteerd en worden genen uitgewisseld. Dit kan door het kiezen van een cross-over point, alle genen voor het cross-over point van het eerste chromosoom worden dan gecombineerd met alle genen achter het cross-over point van het tweede chromosoom en andersom. Een alternatief is om willekeurige genen van het eerste chromosoom te selecteren en deze aan te vullen met de overige genen van het tweede ouder chromosoom. De kans dat cross-over plaats vindt hangt af van de cross-over kans, deze is meestal rond 0.7.

Bij mutatie wordt een willekeurige gen iets aangepast. Als de genen binair zijn wordt het bit omgedraaid, bij continue waarden (voor de gewichten) wordt een kleine random waarde (vaak gegenereerd met een normale verdeling met gemiddelde 0) bij de waarde opgeteld. De kans op mutatie is meestal klein, tussen 0.01 en 0.001.

Selectie, cross-over en mutatie worden herhaald tot de nieuwe populatie even groot is als de oorspronkelijke populatie. Vervolgens wordt deze populatie gebruikt in de volgende iteratie. Deze cyclus wordt een vooraf bepaald aantal stappen herhaald of tot een gewenste fitness is bereikt.

(b) (2 punten) Wat is het voordeel van genetische algoritmen ten opzichte van het backpropagation algoritme dat normaal gebruikt wordt voor het trainen van multilayer neurale netwerken?

Antwoord: Backpropagation is een hill-climbing algoritme dat erg makkelijk in een lokaal optimum (configuratie van de gewichten) blijft steken. Genetische algoritmen zijn beter in staat een globaal optimum te vinden (hoewel er geen garantie is).

15. (5 punten) De particles in *Local best Particle Swarm Optimization* passen hun positie \mathbf{x}_i in de oplossingsruimte aan volgens $\mathbf{x}_i(t+1) = \mathbf{x}_i(t) + \mathbf{v}_i(t+1)$. De elementen van de snelheidsvector $\mathbf{v}_i(t+1)$ worden berekend volgens:

$$v_{ij}(t+1) = v_{ij}(t) + c_1 r_{1j} [y_{ij}(t) - x_{ij}(t)] + c_2 r_{2j} [\hat{y}_{ij}(t) - x_{ij}(t)] \quad (12)$$

Leg uit waar de termen in deze vergelijking voor staan, met andere woorden leg (in woorden) uit hoe de snelheid van een particle wordt bepaald.

Antwoord: De snelheid van een particle hangt af van de vorige snelheid, de afstand en positie van het particle t.o.v. van zijn beste positie (cognitieve component) en zijn afstand en positie t.o.v. de beste positie van de neighborhood (sociale component). c_1 en c_2 zijn constanten die de bijdragen van deze componenten wegen. r_{1j} en r_{2j} zijn random getallen tussen 0 en 1.

Einde tentamenopgaven.

Controleer voor de zekerheid of je alle vragen hebt beantwoord. Het zouden er 15 moeten zijn.