

Tentamen

TI2730-B – Computational Intelligence

IN2205 – Kennissystemen

09-nov-2012, 14:00–17:00

- Dit tentamen heeft 7 meerkeuzevragen in totaal goed voor 14 punten en 6 open vragen met in totaal 36 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:	8	9	10	11	12	13	Totaal:
Punten:	8	5	8	4	4	7	36

- Het gebruik van boek of aantekeningen tijdens dit tentamen is niet toegestaan.
- Het gebruik van een rekenmachine is toegestaan.
- Controleer, voordat je je antwoorden inlevert, 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 Evolutionary Computation, Bayesian Reasoning en Swarm Intelligence, en de slides.
- 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: 4.

Succes!

Meerkeuzevragen

1. (2 punten) We hebben drie fuzzy sets A , B en C met membershipfuncties:

$$\mu_A(x) = (0.1/1, 0.2/2, 0.3/3, 0.4/4, 0.5/5, 0.6/6, 0.7/7) \quad (1)$$

$$\mu_B(x) = (0.8/1, 0.1/2, 0.6/3, 0.3/4, 0.7/5, 0.2/6, 0.9/7) \quad (2)$$

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

Hoe ziet de membershipfunctie van de fuzzy set $(A \cap B) \cup \neg C$ eruit als je deze berekent met behulp van de standaardoperatoren?

- A. $\mu_{(A \cup \neg B) \cap C}(x) = (0.5/1, 0.3/2, 0.4/3, 0.7/5, 0.6/5, 0.2/6, 0.7/5)$
 - B. $\mu_{(A \cup \neg B) \cap C}(x) = (0.2/1, 0.7/2, 0.4/3, 0.3/4, 0.4/5, 0.8/6, 0.5/7)$
 - C. $\mu_{(A \cup \neg B) \cap C}(x) = (0.5/1, 0.7/2, 0.6/3, 0.4/4, 0.4/5, 0.9/6, 0.5/7)$
 - D. $\mu_{(A \cup \neg B) \cap C}(x) = (0.2/1, 0.2/2, 0.5/3, 0.4/4, 0.5/5, 0.5/6, 0.5/7)$
2. (2 punten) Bij het evalueren van Fuzzy regels met behulp van Mamdani inference gaat het proces als volgt:
- A. Crisp input, fuzzyficatie, combineren van regels, evalueren van regels, defuzzificatie van output, crisp output.
 - B. Crisp input, evalueren van regels, fuzzyficatie, combineren van regeloutput, defuzzificatie van output, crisp output.
 - C. Crisp input, fuzzyficatie, evalueren van regels, combineren van regeloutput, defuzzificatie, crisp output.
 - D. Fuzzy input, defuzzificatie van input, evalueren van regels, combineren van regeloutput, fuzzificatie, fuzzy output.
3. (2 punten) Fuzzy logic (FL), Certainty factoren (CF) en Bayesiaanse netwerken (BN) zijn alle drie manieren om met onzekerheid te kunnen redeneren. Welke van de volgende uitspraken is waar:
- A. FL, CF en BN zijn gebaseerd op kansrekening.
 - B. Een CF is een schatting van de zekerheid van een regel, BN is gebaseerd op kansrekening.
 - C. FL is een extensie op binaire set membership, CF en BN zijn gebaseerd op kansrekening.
 - D. Geen van de manieren hebben te maken met kansrekening, BN gaat namelijk over causality.
4. (2 punten) In reinforcement learning representeert de value function $V(s)$...
- A. de door de ontwerper bepaalde positieve en negatieve rewards voor acties leidend tot state s
 - B. het mechanisme om cumulatieve toekomstige reward te leren
 - C. de geleerde utiliteit van een state s
 - D. de door de ontwerper bepaalde utiliteit van een state s
5. (2 punten) Competitive learning is een vorm van:
- A. Unsupervised learning
 - B. Supervised learning
 - C. Semisupervised learning
 - D. Reinforcement learning

6. (2 punten) In Particle Swarm Optimization (PSO) is de snelheid van een deeltje i in dimensie j gegeven door:

$$v_{ij}(t+1) = v_{ij}(t) + c_1 r_{1j}(t) [y_{ij}(t) - x_{ij}(t)] + c_2 r_{2j}(t) [\hat{y}_j(t) - x_{ij}(t)] \quad i = 1, \dots, N$$

Met N het aantal deeltjes. Als de versnellingscoëfficiënt $c_2 \gg c_1$ en $r_{1j} = r_{2j}$ dan geldt dat:

- A. het cognitive component dominant is, wat random exploratie van de ruimte faciliteert
 - B. het sociale component dominant is, wat random exploratie van de ruimte faciliteert
 - C. het cognitive component dominant is, wat het risico op convergentie naar een lokaal minimum voor alle deeltjes vergroot
 - D. het sociale component dominant is, wat het risico op convergentie naar een lokaal minimum voor alle deeltjes vergroot
7. (2 punten) Het 'schema theorem' voor genetische algoritmes:
- A. garandeert convergentie van het algoritme naar een optimale oplossing
 - B. garandeert dat het aantal schemata met een bovengemiddelde fitness zal toenemen in de volgende generatie
 - C. garandeert dat het aantal oplossingen met een bovengemiddelde fitness zal toenemen in de volgende generatie
 - D. garandeert convergentie van het algoritme in een eindig aantal stappen.

Open vragen

8. Geef voor ieder van de onderstaande vraagstukken aan welke van de in dit vak behandelde oplossingsmethoden het meest geschikt is. Motiveer je antwoorden.
- (a) (2 punten) Ik wil een container van 10m³ vullen met dozen van verschillende dimensies, waarbij de container zo vol mogelijk wordt gepakt.
 - (b) (2 punten) Ik wil een systeem bouwen dat aan de hand van webcam input bepaald of er een gezicht in beeld is.
 - (c) (4 punten) Voor rampsituaties in gebouwen wil ik een detectie systeem maken met kleine, snelle robotjes dat bepalen kan waar er nog mensen binnen zijn.

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

- 1. IF f THEN d
- 2. IF c AND b THEN q
- 3. IF q THEN z
- 4. IF h THEN f
- 5. IF a OR b THEN s
- 6. IF d THEN s
- 7. IF b AND z THEN w
- 8. IF g THEN h
- 9. IF b THEN h
- 10. IF d AND z THEN x

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

- (a) (3 punten) Laat zien hoe de inference engine aantoont of doel x waar is. Geef aan welke regels er gebruikt worden, welke doelen er op de stack geplaatst worden en hoe de feiten database verandert.
- (b) (2 punten) In dezelfde knowledge base: stel dat regel 10 een certainty factor van 0.5 heeft, en regel 1 een certainty factor van 0.9. Wat is dan de zekerheid over conclusie x , gegeven dat zowel f als z een certainty factor van 0.8 hebben?

10. Ik kwam van de week thuis en had 's morgens een was ingezet. Bij thuiskomst was de vloer bij de wasmachine helemaal nat. Omdat ik net een nieuwe wasmachine heb, wil ik graag weten of de oorzaak van de plas water komt door een kapotte wasmachine of gewoon omdat ik het putje weer eens moet onstoppen. Als de wasmachine stuk is, dan wordt de was meestal ook niet schoon. Ik schat de kans dat de was schoon is als de wasmachine kapot is op ongeveer 0.1, terwijl de kans op schone was bij een goed werkende wasmachine ongeveer 0.7 is. Verder denk ik dus dat er twee mogelijke oorzaken zijn voor de plas water: een verstopt putje en een kapotte wasmachine. De kans dat er een plas water ligt als het putje niet verstopt is en de wasmachine wel goed werkt is 0. Dat gebeurt nooit. De kans dat er een plas water ligt als het putje verstopt zit is groot, 0.9, en het maakt niet uit of de wasmachine nu wel of niet stuk is. Als alleen de wasmachine stuk is, en het putje niet verstopt, dan schat ik de kans dat er water ligt op ongeveer 0.3. Als het putje verstopt zit, dan hoor ik meestal ook geborrel, die kans schat ik in op 0.8. Als het putje niet verstopt zit, kan er nog steeds wel geborrel te horen zijn, maar de kans daarop is beduidend kleiner, ongeveer 0.2.
- (a) (2 punten) Teken, aan hand van bovenstaande gegevens, een Bayesiaans netwerk waarmee we kunnen bepalen wat de oorzaak van de overstroming is. Wat zijn de variabelen, wat zijn de conditionele kansen, en hoe lopen de afhankelijkheden?
 - (b) (4 punten) Bereken de kans dat de wasmachine kapot is, en laat de berekening zien.
 - (c) (2 punten) Omdat ik zeker wil weten dat de wasmachine niet stuk is, luister ik goed naar de afvoer tijdens het pompen. Ik hoor inderdaad geborrel. Wat betekent dit voor de kans dat de wasmachine kapot is, wordt deze groter, kleiner of blijft die gelijk? Leg uit aan de hand van het bayesiaanse netwerk. Je hoeft niets uit te rekenen, maar de uitleg moet wel volledig zijn en kloppen met het netwerk.
11. Multilayer feedforward neural networks zijn de meest gebruikte neurale netwerken. Om output gewichten te updaten moet de volgende term uitgerekend worden:

$$\Delta w_{jk}(p) = \alpha y_j(p) \delta_k(p) \quad (4)$$

- (a) (1 punt) Leg uit wat α is en hoe deze term het leren beïnvloedt.
 - (b) (1 punt) Leg uit wat $y_j(p)$ is en hoe deze functie gekozen moet worden
 - (c) (2 punten) Leg uit wat $\delta_k(p)$ is en hoe deze afgeleid kan worden (in je uitleg hoeft je geen formules te gebruiken)
12. Deze vraag gaat over 'associative memories'.
- (a) (1 punt) Wat is een associative memory en waar wordt het voor gebruikt?
 - (b) (1 punt) Beschrijf de structuur van een Hopfield netwerk (aantal inputs, outputs, layers, weights) en geef de activatie functie voor de neuronen?
 - (c) (1 punt) Wat zijn de belangrijkste beperkingen van een Hopfield netwerk?
 - (d) (1 punt) Wat zijn de belangrijkste verschillen tussen een Hopfield netwerk en een Bidirectional associative memory (BAM)?
13. Een onderzoeker wil een nieuw algoritme ontwikkelen voor het verbeteren van de beeldkwaliteit van plaatjes. Op een plaatje zijn 6 operaties mogelijk (histogram equalization, contrast stretching, sharpening, denoising, color balancement correction, color saturation increase) die achter elkaar uitgevoerd moeten worden. Elke operatie kan op een combinatie van elk van de drie kleurkanalen van het plaatje (R, G, B) uitgevoerd worden. Bijvoorbeeld sharpening op R en G maar niet op B. De volgorde van operaties en de kleurkanalen waarop ze uitgevoerd worden zijn beiden van belang voor de uiteindelijke kwaliteit. Elke operatie kan slechts een keer gebruikt worden. Het algoritme bestaat dus uit een volgorde van 6 operaties waarbij voor elke operatie aangegeven moet worden op welk(e) kleurkanaal(en) deze uitgevoerd moet worden.
- De onderzoeker heeft een testprogramma waarmee de beeldkwaliteit van een plaatje getest kan worden. Dit programma geeft voor een plaatje een getal tussen [0,1], waarbij 1 hoge kwaliteit is, en 0 lage. Omdat de onderzoeker een algoritme wil om de beeldkwaliteit te verhogen dat voor alle mogelijk plaatjes werkt, heeft de onderzoeker ook een set van 300 plaatjes om het nieuwe algoritme op uit te proberen.

Leg uit hoe de onderzoeker met behulp van een genetisch algoritme het nieuwe algoritme kan ontwikkelen. Het algoritme moet zodanig werken dat het gemiddeld genomen de beste kwaliteitsverbetering laat zien.

- (a) (1 punt) de codering als chromosoom (het genotype),
- (b) (1 punt) de populatie,
- (c) (1 punt) de fitnessfunctie,
- (d) (1 punt) selectie,
- (e) (1 punt) cross-over,
- (f) (1 punt) mutatie,
- (g) (1 punt) de nieuwe populatie en terminatie

Einde tentamenopgaven.

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