

Tentamen TI2736-A – Computational Intelligence

5-nov-2014, 14:00–16:30

- Het hele tentamen heeft 5 meerkeuzevragen in totaal goed voor 10 punten en 5 open vragen met in totaal 37 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.
- 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 de onderwerpen zoals behandeld tijdens de colleges.
- 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) Competitive learning is a form of:
 - A. reinforcement learning.
 - B. supervised learning.
 - C. unsupervised learning.
 - D. swarm intelligence.
2. (2 punten) Local-best and Global-best PSO differ in:
 - A. The formulation of the particles velocity.
 - B. The computation of the Social Component.
 - C. The value of the fitness function
 - D. The computation of the cognitive Component.
3. (2 punten) Which of these activation functions cannot be used in a feed forward neural network, if we plan to train it with backpropagation:
 - A. A hyperbolic tangent
 - B. A sigmoid.
 - C. A sinusoid
 - D. A step function
4. (2 punten) Bij Reinforcement Learning (RL) wordt er geleerd om acties in states wel of niet te herhalen op basis van het verkrijgen van reinforcement (feedback) na het uitvoeren van een actie. Hierbij vergroot positieve feedback de kans op herhaling, terwijl negatieve feedback deze kans verkleint. Welke van de volgende stellingen is onwaar:
 - A. $V(s)$ de discount factor γ zorgt er voor dat naar mate reinforcement verder in de toekomst verwijderd is deze minder waard is voor de berekening van de waarde van een state s .
 - B. Het aantal mogelijke acties en het aantal mogelijke states bepaalt in sterke mate de leerbaarheid van een optimale oplossing.
 - C. Reinforcement is positief of negatief en wordt formeel genoteerd als $R(s, a, s')$.
 - D. Omdat RL een Computational Intelligence techniek is, is het erg goed in het leren van optimaal gedrag in hele grote state spaces.
5. (2 punten) In Computational Intelligence is de fitness functie wiskundig gezien:
 - A. Een projectie van states naar waardes.
 - B. Een landschap met bergen en dalen.
 - C. Een manier om zoeken te optimaliseren.
 - D. Een projectie van waardes naar states.

Open vragen

6. De onderstaande vraagstukken gaan over CI technieken in de praktijk. Geef voor ieder van de onderstaande vraagstukken aan welke van de *in dit vak* behandelde oplossingsmethoden het *meest* geschikt is (zijn). Motiveer je antwoorden door kort(!) te beschrijven (a) wat voor TYPE problemen (beslisprobleem, optimalisatieprobleem, etc..) een rol in het vraagstuk spelen, (b) waarom juist jouw gekozen methodes geschikt zijn voor deze problemen, en (c) hoe je met die methodes de problemen oplost (het laatste mag kort, maar moet wel inzicht verschaffen in waarom de methode het probleem op kan lossen).
 - (a) (3 punten) Een afgevaardigde van de NASA komt bij je langs met het volgende probleem. Ze willen een wereldwijd netwerk van satellieten die een vernieuwde versie van GPS moeten opleveren. Kern is

- dat ze de n satellieten zo in een baan om de aarde willen schieten dat dit (a) in drukk bewoonde gebieden een fijnmazig netwerk geeft, en (b) overal op aarde tenminste minimale dekking geeft (of te wel, overal op aarde moet tenminste 3 satellieten meetbaar zijn). Een satelliet kan boven een willekeurig te kiezen breedte en lengtegraad (de locatie) op aarde blijven zweven (ze zijn *geostationair*). Hoe los je dit op?
- (b) (3 punten) Om de veiligheid op de weg te garanderen wil rijkswaterstaat een systeem dat voorspelt hoe druk de snelweg wordt als er een rijbaan dichtgezet wordt. Alle snelwegen in Nederland zijn ingedeeld in korte segmenten. Het wegennet is dus een groot netwerk van segmenten. Elk segment heeft een camera die naar een centraal computer systeem loopt. Om te beginnen willen ze precies weten hoeveel autos er op een segment rijden. Hoe los je dit op? Daarna willen ze voorspellen of er file ontstaat (ja/nee) door het dichtzetten van een rijbaan op een segment. De voorspelling moet op basis van de drukte op dat segment alsmede de 5 segmenten ervoor en 5 erna zijn. Hoe los je dat op?
7. Binnen Computational Intelligence kun je stellen dat exploratie staat voor "zoeken" terwijl exploitatie staat voor "het uitbuiten van een oplossing met hoge fitness". Leg uit, door gebruik te maken van parameters in de formules, hoe exploratie en exploitatie tot uiting komen in....
- (a) (2 punten) ...de Boltmann (soft-max) actie selectie (bijv. bij Reinforcement Learning).
- (b) (2 punten) ...de standaard formule voor een ACO (Ant Colony Optimization).
- (c) (2 punten) ...het standaard protocol voor een Evolutionary Algorithm.
- (d) (2 punten) ...de standaard formule voor PSO (Particle Swarm Optimization).
- (e) (2 punten) Leg ten slotte in algemene bewoording uit wat de exploratie-exploitatie tradeoff is, en geef hierbij aan waarom het een tradeoff is, met andere woorden, wat er gebeurt bij te veel of te weinig van het een of het ander.
8. Ik kwam van de week thuis en had 's morgens een was ingezet. Bij thuiskomst was de vloer bij de wasmachine weer helemaal nat. Omdat mijn wasmachine nog steeds best nieuw is, 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. De kans op een kapotte wasmachine is in het algemeen klein, zo'n 0.05, terwijl de kans op een verstopt putje groter is, zo'n 0.1. 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 0.9 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 0.01. Dat gebeurt bijna nooit. De kans dat er een plas water ligt als het putje verstopt zit is groot, 0.8, en dat is niet afhankelijk van 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 0.7. 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) (3 punten) Bereken de kans dat de wasmachine kapot is, gegeven dat we weten dat de vloer nat is. Ik heb nog niet aan de was geroken su ik weet verder niets over of de was schoon, en omdat ik er niet bij was tijdens het wassen weet ik ok niet of er geborrel te horen was. Laat je berekening zien.
- (c) (2 punten) Als ik de volgende keer de wasmachine aandoe, dan luister ik eens goed wat er gebeurt. Ik hoor geen 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, precies kloppen met de afhankelijkheden in het netwerk, en je moet gebruik maken van de variabelen en of ze true of false zijn/worden.

9. Consider a Hopfield network where you want to store the fundamental memories:

$$\mathbf{x}_1 = \begin{bmatrix} -1 \\ -1 \\ 1 \\ 1 \end{bmatrix}; \mathbf{x}_2 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix} \quad (1)$$

- (a) (1 punt) design the Hopfield network that can accomplish the task. Indicate the number of input and output neurons and the number of layers. Give a graphical representation of the network, specifying where the input signals enter the network and how the neurons are connected.
- (b) (2 punten) remembering that the learning rule for a Hopfield network that needs to store M fundamental memories is:

$$W = \sum_{m=1}^M \mathbf{x}_m \mathbf{x}_m^T \quad (2)$$

train the network you designed at point (a) to store the fundamental memories $\mathbf{x}_1$ and $\mathbf{x}_2$.

- (c) (1 punt) Use the trained network to retrieve the fundamental memories $\mathbf{x}_1$ and $\mathbf{x}_2$ remembering that

$$Y = \text{sign}(W\mathbf{x}_m + \Theta) \quad \text{and} \quad \Theta = \mathbf{0} \quad (3)$$

- (d) (2 punten) what is a main application for Hopfield networks? Show whether your trained network can be used to do the job (hint: flip a bit of $\mathbf{x}_2$ and test your network on it).
- (e) (1 punt) what is the main limitation of Hopfield networks?

10. A care home needs to plan in advance the holiday schedule of its employees for the Christmas break, which lasts for 3 weeks around Christmas. There are 8 nurses working in the care home, each nurse works 5 days a week, and each of them wants to go on holidays for either one or two weeks. If a nurse goes on holidays for 2 weeks, these two have to be consecutive. The main concern of the care home is that there should always be a 24 hour coverage every day. Although each nurse works 5 days a week, they work in shifts and their shifts are not of equal length. Each nurse can cover at most one shift per day, and there is no need to have overlap between the nurses shifts (that is, when one starts the shift, the other can finish). So, what the care home wants, is to schedule the holidays of the nurses so that every week the remaining nurses can cover with their shifts at least 24 hours per day. Another way to put it is that the sum of the length of the shifts per day of the nurses that are not on holiday should be equal or greater than 24 (and more is better to cover emergencies). Knowing that the nurses have different working shift lengths, and request the amount of holidays indicated in the table below:

nurse	Number of weeks of holidays requested	Number of worked hours per shift
1	1	4
2	1	8
3	2	8
4	2	4
5	1	4
6	2	8
7	1	8
8	1	6

Use an evolutionary computation method to find the schedule that (1) allows all nurses to go on holidays and (2) maximises the time covered by the shifts of the available nurses In particular:

- (a) (1 punt) Indicate whether you would use a Genetic Algorithm or an Evolutionary strategy to do the job, and motivate your answer.
- (b) (1 punt) Describe how you would encode your candidate solutions into a chromosome, also specifying

what each gene represents. Schemes are welcome. Please justify your choices.

- (c) (1 punt) Describe how you would initialize the population
- (d) (1 punt) Define the fitness function
- (e) (1 punt) Describe your selection scheme
- (f) (1 punt) Describe genetic operators, taking care of adapting their functioning to this specific problem, if necessary
- (g) (1 punt) Indicate when/why the algorithm will terminate.

Einde tentamenopgaven.

Controleer voor de zekerheid of je alle vragen hebt beantwoord. Het zouden er 10 moeten zijn.