

Naam:

Collegekaart-nummer:

Versie opgavenblad (omcirkelen): A | B | C | D

Dit tentamen duurt 3 uur. Het is verboden literatuur, aantekeningen, een programmeerbare rekenmachine, of een telefoon te gebruiken. Het gebruik van een standalone niet-programmeerbare rekenmachine is toegestaan.

Open vragen

Antwoorden op alle vragen worden ingevuld op deze bladen. Vaak is het handig je antwoord op klad voor te bereiden. Twee punten voor elk sub-onderdeel, tenzij anders aangegeven. Antwoorden op open vragen zonder berekening of motivatie leveren geen punten op.

1. (Co-evolutie, Watson & Pollack.) We beschouwen de *Min-d* fitness-maat in twee dimensies. Illustreer met een getallen-voorbeeldje in \mathbb{R}^2 de volgende eigenschappen:

(a) Eigenschap 1: *Min-d* is intransitief. (Tip: probeer eerst een aantal gevalletjes uit op kladpapier.)

Antwoord. We moeten puntenparen (a, b) , (c, d) , (e, f) vinden, zo dat $(a, b) \leq_{Min-d} (c, d)$, $(c, d) \leq_{Min-d} (e, f)$, maar niet $(e, f) \leq_{Min-d} (a, b)$. Dat lukt bijvoorbeeld met $(1, 1)$, $(2, 3)$, en $(3, 0)$. Een dergelijk tegenvoorbeeld kan worden gevonden door uitproberen. Kies de eerste twee punten willekeurig, en stuur de keuze van het derde punt (e, f) zó dat deze in de oorspronkelijke dimensie (hier x) groter is dan (c, d) maar ten opzichte van (a, b) in een andere dimensie vergeleken moet worden, en dan kleiner is.

(b) Eigenschap 2: *Min-d* is niet-monotoon stijgend op attribuutwaarden. D.w.z.: als $x \leq_{Min-d} y$, en y' ontstaat uit y door de waarde van één van haar attributen te verhogen, dan is het niet noodzakelijk zo dat $x \leq_{Min-d} y'$.

Antwoord. We moeten puntenparen (a, b) , (c, d) , (e, f) vinden, zo dat $(a, b) \leq_{Min-d} (c, d)$, $c < e$ en $d \leq f$ (of $c \leq e$ en $d < f$), maar niet $(a, b) \leq_{Min-d} (e, f)$. Dat lukt bijvoorbeeld met $(2, 2)$, $(3, 0)$, en $(4, 1)$. Een dergelijk tegenvoorbeeld kan ook weer worden gevonden door uitproberen.

2. In Fig. 1 is de uitbetalingsmatrix gegeven van een twee-persoons competitief symmetrisch niet-nulsom spel op basis van volledige informatie met simultane zetten en kwantitatieve beloningen.

(a) Bepaal het Pareto front.

Antwoord. Het uitbetalingsprofiel $\{(3, 4)\}$ domineert alle andere uitbetalingsprofielen.

(b) Bepaal alle pure Nash evenwichten.

Antwoord. Pure Nash evenwichten: $\{(0, 0), (1, 1)\}$. Dus beiden spelen hun eerste actie of beiden spelen hun tweede actie. In deze scenario's hebben beiden partijen geen behoefte om hun actie te wijzigen.

(c) Laat $p, q \in [0, 1]$ de gemixte strategie van A resp. B zijn. Bepaal $\partial \text{Payoff}_A / \partial p$ en $\partial \text{Payoff}_B / \partial q$.

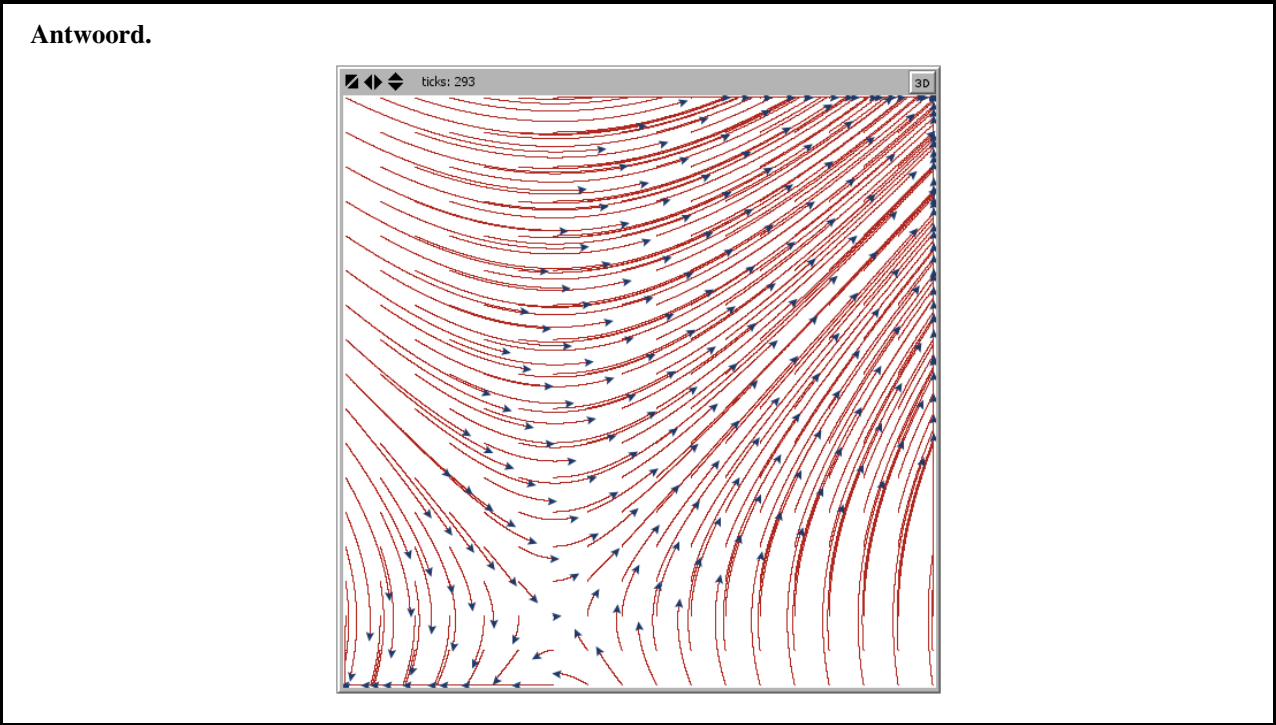
Antwoord.

$$\begin{cases} \partial \text{Payoff}_A / \partial p = 9q - 1 \\ \partial \text{Payoff}_B / \partial q = 11p - 4 \end{cases}$$

(d) Bepaal alle pure en gemixte Nash evenwichten.

Antwoord. Voor gemixt evenwicht afgeleiden payoffs nul stellen en oplossen. Nash evenwichten: $\{(0, 0), (4/11, 1/9), (1, 1)\}$.

(e) Teken het bijbehorende gradient-veld. Besteed aandacht aan hoeken, randen, en bijzondere punten.



(f) Stel, de “3” linksboven wordt vervangen door een x , waarbij x een reëel getal is. Bepaal voor dit algemenere geval het Pareto front.

Antwoord. De uitbetalingsprofielen $(-1, -3)$ en $(-5, -2)$ worden gedomineerd door $(0, 2)$, dus die vallen af. Het uitbetalingsprofiel $(0, 2)$ domineert $(x, 4)$ nooit, daar zorgt de 4 voor. Het uitbetalingsprofiel $(0, 2)$ wordt gedomineerd door $(x, 4)$ als en slechts als $x > 0$.
 Het Pareto-front is nu dus gelijk aan:

$$\begin{cases} \{(x, 4), (0, 2)\} & \text{als } x \leq 0, \\ \{(x, 4)\} & \text{anders.} \end{cases}$$

3. In Fig. 2 is een toestandengraaf voor een reinforcement leer-probleem afgebeeld. Discount $\gamma = 1$. Stel, de volgende vier runs worden achtereenvolgens afgelegd:

- 1: C, B, D, T2
- 2: A, B, D, T1
- 3: C, B, D, T1
- 4: C, B, D, T2

(a) Alle V -waarden zijn geïnitieerd op nul. Laat met een tabel zien hoe V -waarden van toestanden worden geupdate bij Monte-carlo leren.

Antwoord.

Monte-Carlo. Verdiscontering gamma = 1

Epoch	A	B	C	D	T1	T2
C, B, D, T2		6	-1	3		0
A, B, D, T1	5	7		4	0	
C, B, D, T1		7	0	4	0	

C, B, D, T2	6	-1	3	0	
First visit gem.	5	6.5	-0.7	3.5	0 0
Every visit gem.	5	6.5	-0.7	3.5	0 0

- (b) Alle V -waarden zijn geïnitieerd op nul. Laat met een tabel zien hoe V -waarden van toestanden worden geupdate bij temporal difference leren met leerfactor $\alpha = 0.5$.

Antwoord.

Temporal Difference. gamma = 1; leerfactor alpha = 0.5

Epoch	A	B	C	D	T1	T2
C, B, D, T2		1.5	-3.5	1.5		0
A, B, D, T1	-0.3	3		2.8	0	
C, B, D, T1		4.4	-3.8	3.4	0	
C, B, D, T2		5.4	-3.2	3.2		0
Laatste waarde	-0.3	5.4	-3.2	3.2	0	0

4. (Vier punten.) Bespreek de Gaia hypothese. Formuleer je bespreking eerst in klad. Schrijf het daarna over in onderstaande box.

Er wordt gelet op het volgende. Dit lijstje is bedoeld om enige duidelijkheid te verschaffen—niet om te intimideren: je bespreking wordt met welwillendheid nagekeken.

- i) Taalgebruik, formulering, en opbouw.
- ii) Samenhang en betekenis.
- iii) Het (zinvol) voorkomen van essentiële begrippen en/of categorieën.
- iv) Lengte. (Nodeloos lange antwoorden krijgen puntenaftrek.)

Antwoord.

James Lovelock (1919) was vanaf 1965 betrokken bij onderzoek van de NASA naar buitenaards leven. De vraag was toen of er (primitief) leven op Mars is. Lovelock's antwoord was: kijk naar de samenstelling van de atmosfeer. Geen leven impliceert chemisch evenwicht (zie Mars), dus geen chemisch evenwicht (Aarde) impliceert leven. Leven = respiratie (ademhaling), metabolisme (stofwisseling), zelf-regulering, feedback, prikkelbaarheid, voorplantingsvermogen. Op basis van deze observaties bedacht Lovelock de Gaia hypothese: behalve de organismen die op aarde leven, leeft de aarde ook, en oefent terug invloed uit op de organismen. Leven creëert zijn eigen ideale omstandigheden door de aardse omgeving te beïnvloeden, en omgekeerd.

In samenhang met de Gaia hypothese presenteerde Lovelock in 1983 het Daisyworld model (Daisy = madeliefje). Daisyworld is een computermodel van een denkbeeldige planeet waar witte en zwarte daisies leven. Daisies kunnen groeien, reproduceren, hun omgeving veranderen en sterven. Het doel van het Daisyworld model was het opzetten van een kunstmatige omgeving waarin de wisselwerking tussen organismen en hun omgeving kon worden verkend.

Er is een zwakke Gaia en een sterke Gaia:

- Zwak: wisselwerking is metafoor om processen te begrijpen. Pytho-plankton (zee-algen) die di-methylsulfide (DMS) uitscheiden illustreren de (zwakke) Gaia-hypothese.
- Sterk: "het" is er echt: de biosfeer wordt doelbewust in stand gehouden door een superorganisme, de aarde. Acceptatie:
 - Zwak: geen probleem.
 - Sterk: pseudo-wetenschappelijk, niet empirisch te bewijzen en in strijd met de goed onderbouwde evolutietheorie.

(Dictaat Wiering 5.3.2: Gaia-hypothese; slides Gaia & Co-evolutie: Gaia-hypothese.)

Meerkeuze vragen

Bij elke meerkeuzevraag is steeds precies één antwoord het juiste. Twee punten per onderdeel.

- Het principe van de Rode Koningin behelst het volgende.
 - Bij een fitness-proportionele selectie dient in ieder geval het meest fitte individu meegenomen te worden naar de volgende generatie.
 - In co-evolutie moeten individuen blijven verbeteren om fit te blijven.
 - Het co-evolutieproces verloopt sneller als individuen vaker gekruist worden met fitte individuen.
 - Exploratie op feromoon wordt bepaald door de laatste winnaar.
- In "Coevolutionary dynamics in a minimal substrate" werken Watson en Pollack met individuen waarvan de fitness niet wordt gerepresenteerd door één getal maar door een 2-dimensionale vector. Later worden zelfs 10-dimensionale vectoren gebruikt. Waarom is dit, althans volgens de auteurs, zinvol?
 - Het succes van een individu kan niet altijd in één dimensie worden afgemeten en hangt typisch af van meerdere factoren.
 - Omdat van vectoren de norm kan worden bepaald.
 - Op deze manier kunnen dimensies bij elkaar worden opgeteld om de absolute fitness uit te rekenen.
 - Het goede antwoord staat er niet bij.
- Watson en Pollack kiezen er in hun experiment voor om genotypen te representeren door middel van bitstrings. Wat is, wat hen betreft, hiervoor het meest belangrijke motief?
 - Op deze manier laten bitstrings zich op een voor de hand liggende manier kruisen.
 - Op deze manier laten bitstrings zich op een voor de hand liggende manier muteren.
 - Op deze manier zijn mutaties evenredig verdeeld over bitstrings.
 - Op deze manier is bij een fit individu de kans op verslechtering bij muteren aanmerkelijk groter is dan de kans op verbetering.

Antwoord. Het goede antwoord is in overeenstemming met de genetica, meer i.h.b. met wat de gezaghebbende evolutiebioloog R.A. Fischer heeft gezegd, namelijk dat in latere stadia van de evolutie mutaties bijna altijd een nadelig effect hebben (R.A. Fischer, The genetical theory of natural selection, 1930).

- Gegeven zijn twee populaties A en B met absolute fitness

A :	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9
Fitness :	3	2	8	9	1	1	3	5	6
B :	b_1	b_2	b_3	b_4	b_5	b_6	b_7	b_8	b_9
Fitness :	3	5	2	7	2	1	6	9	4

De referentie-set, S , is gelijk aan heel B . Bereken de genormaliseerde relatieve fitness van a_9 .

- 6/9
- 6/34
- 6/43
- 9/43

Antwoord. Relatieve fitness van a_1 t/m a_9 : 3, 1, 8, 8, 0, 0, 3, 5, 6. Genormeerde relatieve fitness: 3/9, 1/9, 8/9, 8/9, 0/9, 0/9, 3/9, 5/9, 6/9. Totaal: 34/9. Genormaliseerde relatieve fitness: 3/34, 1/34, 8/34, 8/34, 0/34, 0/34, 3/34, 5/34, 6/34.

- Na het verschijnen van het boek "Perceptrons" (Minsky en Papert, 1969), zakte het onderzoek naar Neurale netwerken in. Waarom was dat?
 - Minsky en Papert lieten zien dat het lineaire perceptron zelfs de meest eenvoudige logische schakelingen niet kan representeren.
 - Minsky en Papert bewezen dat de schakeltijd van het lineaire perceptron te traag is om er intelligente netwerken van te kunnen bouwen.
 - Door middel van een gegeneraliseerd tel-argument beargumenteerden Minsky en Papert dat kunstmatige neurale netwerken nooit intelligent gedrag kunnen opleveren.
 - Het goede antwoord staat er niet bij.
- Waar word de hypotheseruimte in een neuraal netwerk door vertegenwoordigt?
 - Door de gewichten.
 - Door de netwerk-topologie, d.w.z., de knopen en de links.
 - Door de topologie en de gewichten.
 - Door de topologie, en de respectievelijke leervoorbeelden.

Antwoord. Alleen de gewichten, de rest ligt tijdens het leren vast.

- Welke van de volgende update-methoden zijn gevoelig voor de volgorde waarin data $\{d_k\}_{k=1}^{\infty}$ wordt aangeboden?

i) Standaard gemiddelde update:

$$x = x + k^{-1}(d_k - x)$$

ii) Standaard geometrische update:

$$x = x + \alpha(d_k - x), \text{ waarbij leerfactor } 0 \leq \alpha \leq 1.$$

- (a) *i*.
 ✓ *ii*.
 (c) *i* en *ii*.
 (d) Geen.

8. We leren voorbeelden in R^3 met een standaard twee-dimensionaal Kohonen-netwerk. Het netwerk is random geïnitieerd binnen het bolletje om $(0, 0, 0)^T$ met straal 0.01. De leervoorbeelden 1 – 1,000 bevinden zich uniform verdeeld over het oppervlak S van de eenheidsbol B . De leervoorbeelden 1,001 – 20,000 bevinden zich uniform verdeeld over het oppervlak S' van de eenheidsbol B' die ontstaat door B te transleren over de vector $(5, 0, 0)^T$. De leerfactor is 0.4. We kunnen het volgende zeggen over het resulterende netwerk.

- ✓ Het spant zich om B' heen.
 (b) Het spant zich, deels om B , en deels om B' heen.
 (c) Het bevindt zich tussen B en B' .
 (d) Het spant zich deels over het half-rond van B' dat naar B wijst, en deels over het half-rond van B dat naar B' wijst.

Antwoord. 95% extra iteraties om B' heen zorgt er wel voor dat het hele net om B' heen gaat zitten. B wordt vergeten.

9. Welke uitspraken zijn waar?

- i) K -means is batch.
 ii) K -means kan omgaan met ontbrekende attribuut-waarden.
 (a) Geen.
 ✓ Alleen *i*.
 (c) Alleen *ii*.
 (d) *i* en *ii*.

10. Wat is “stochastic approximation to gradient descent”?

- ✓ Online leren van een neurale netwerk. Dit gaat niet strict volgens de fout-functie.
 (b) Online leren van een neurale netwerk. Of je online blijft hangt er maar van af.
 (c) Offline leren van een neurale netwerk. Dit gaat met random gegenereerde voorbeelden.
 (d) Offline leren van een neurale netwerk. Dit gaat met random geïnitieerde gewichten.

11. Het geïtereerde prisoner’s dilemma met payoffs $R = 3$, $T = 5$, $S = 0$, $P = 1$. We laten RAND (samenwerken met kans 0.5) in een gelijke verhouding spelen tegen ALL-C, ALL-D, TFT, en zichzelf. Wat is verwachte gemiddelde opbrengst van RAND?

- (a) 1.5
 ✓ 2.25
 ✓ 2.3
 (d) 3

Antwoord. $(4 + 0.5 + 2.25 + 2.25)/4 = 2.25$. (Het getal 2.3 is een raar alternatief, want 2.25 afgerond tot 1 decimaal is ook 2.3 ...)

12. In het geïtereerde prisoner’s dilemma is TFT is nooit echt veel slechter af dan de tegenstander. Waarom is dat?

- ✓ Met TFT kan de tegenstander verzaken, maar moet bij wisseling van strategie altijd zijn (haar) winst weer inleveren.
 (b) Met TFT kan de tegenstander worden verleid tot samenwerken. (Wat een verstandige tegenstander vervolgens ook zal doen.)
 (c) Met TFT kan de protagonist zelf bepalen wat deze de volgende ronde speelt.
 (d) Met TFT wordt de tegenstander genadeloos geïmiteerd. Dit wekt in het algemeen irritatie en vervolgens misstappen op.

13. We passen kwalitatief gesuperviseerd leren met vector-kwantisatie toe op drie cluster-representanten $A = (1, 77, 1)$, $B = (2, 77, 3)$ en $C = (3, 77, 0)$. Bepaal de bewegingen van deze cluster-representanten als het leervoorbeeld $l = (2, 77, 2) \rightarrow A$ wordt aangeboden.

- (a) Er gebeurt niets.
 (b) Repr- A beweegt naar l toe.
 ✓ Repr- A beweegt naar l toe en Repr- B beweegt van l af.
 (d) Repr- A beweegt naar l toe en Repr- B en Repr- C bewegen van l af.

Antwoord. Letterlijke definitie van het algoritme. Merk op dat de tweede dimensie $y = 77$ kan worden weggegooid, zodat je het voorbeeld in R^2 kunt tekenen.

14. We beschouwen een Markov beslisprobleem waarvan het hele model bekend is. Er zijn tenminste twee manieren om een vaste policy, π , te evalueren.

- (a) Monte-carlo sampling en temporal difference leren.
 ✓ Gauss-eliminatie en policy evaluation.
 (c) Gauss-eliminatie en value-iteration.
 (d) Het goede antwoord staat er niet bij.

Einde van alle opgaven. Heb je je antwoorden gecontroleerd? Vergeet niet de onderwijsenquête in te vullen. Bedankt voor je deelname en een prettige vakantie.

Figuren

		Speler B	
		b_2	b_3
Speler A	a_2	$(3, 4)$	$(-1, -3)$
	a_3	$(-5, -2)$	$(0, 2)$

Fig. 1: Uitbetalingsmatrix van een twee-persoons competitief symmetrisch niet-nulsom spel.

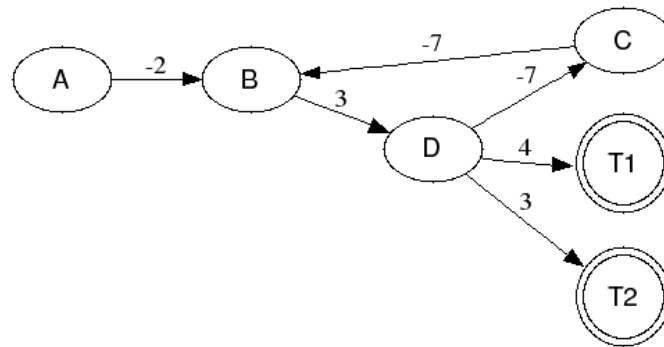


Fig. 2: Toestandengraaf voor een reinforcement leer-probleem.