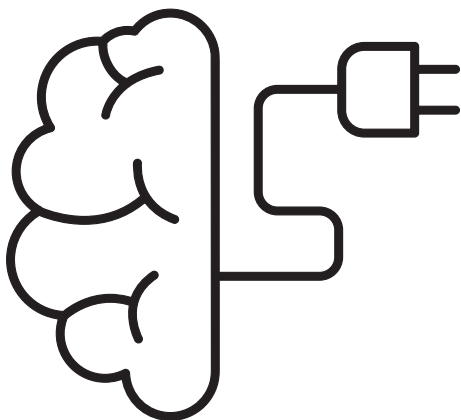


ARTIFICIAL INTELLIGENCE

Op weg naar de artificiële mens?



*Aan Caroline,
voor alles wat echt
en niet artificieel is*

INHOUD

INLEIDING: WAAROM HOUDEN WE ZO VAN ROBOTS	7
ECHTE INTELLIGENTIE EN ARTIFICIËLE INTELLIGENTIE	13
Wat is intelligentie?	13
Artificiële neuronen	17
Het eerste neurale netwerk	19
Backpropagation	20
Feedforward-, recurrente en convolutionele neurale netwerken	21
NIET PROGRAMMEREN, MAAR LEREN	27
Machine learning en deep learning	27
Regressie en classificatie	28
Gesuperviseerd en niet-gesuperviseerd leren	30
Beloningsgebaseerd leren	32
ROBOTS: EEN FYSIEK LICHAAM VOOR AI	37
Belichaamde cognitie en zwermintelligentie	37
Een aangepast lichaam	39
Robots met een fysieke en sociale taak	40
Het uiterlijk van robots	45
EMOTIE	49
Emotie als vorm van intelligentie	49
Emotie laten zien	50
Emotie aflezen	54

TAAL	57
Het belang van taal	57
Spraakherkenning	59
Spraaksynthese	62
Automatisch en statistisch vertalen	63
Neuraal vertalen	66
BEWUSTZIJN	69
De Chinese kamer	69
Soorten bewustzijn	72
Rechten en plichten	75
MACHINES OF IETS MEER?	79
Ethische problemen	79
AI bij de politie en in de rechtbank	82
Moeten we ons dan echt zorgen maken?	85
AANBEVOLEN LECTUUR	88

INLEIDING: WAAROM HOUDEN WE ZO VAN ROBOTS?

Machines fascineren ons. Niet enkel omdat ze ons werk verlichten of omdat ze technisch vernuftig in elkaar steken, maar vaak omdat ze iets menselijks over zich hebben. Een stoommachine lijkt te ademen, een oliepomp heeft net een hoofd dat op en neer beweegt (we noemen die dingen niet voor niks jaknikkers), en de koplampen van auto's zien eruit als twee ogen die de weg afspeuren. We bouwen al duizenden jaren machines om onze fysieke en mentale mogelijkheden uit te breiden, en we zijn al even lang gefascineerd door het bouwen van machines die ons fysiek en mentaal evenbeeld zijn.

Maar het bouwen van een artificiële mens lijkt een prijs te hebben, althans in verhalen. In het verhaal van Frankenstein bijvoorbeeld bouwt de jonge Victor Frankenstein een menselijk wezen, dat hij met bliksem tot leven wekt. Maar Frankenstein walgt van zijn creatie: in plaats van een mens heeft hij een monster gemaakt. Het monster, teleurgesteld door zijn onvolmaaktheid, keert zich tegen zijn schepper en laat een spoor van dood en vernieling na. Toen Mary Shelley *Frankenstein* meer dan 200 jaar geleden schreef, werd elektriciteit nog helemaal niet gebruikt om machines aan te drijven, laat staan om monsters tot leven te wekken – daarvoor was het wachten op de eerste draaiende elektrische motor van Michael Faraday. Maar onze drang om een artificiële mens te bouwen bestaat al vele millennia: in verhalen maakt men

kunstmatige mensen uit klei, ijzer, stoom, lichaamsonderdelen en sinds de vorige eeuw ook uit elektrische onderdelen en elektrische motors. Het woord 'robot' werd voor de eerste keer gebruikt in 1920, in het toneelstuk *Rossumovi Univerzální Roboti* ('Rossums universele robots') van de Tsjechische schrijver Karl Čapek. Sindsdien is de robot als personage niet meer weg te denken uit sciencefictionverhalen. Maar waarom zijn we zo gefascineerd door metalen mensen?

Een eerste verklaring is te zoeken in onze hersenen. Onze hersenen, en die van alle andere diersoorten, zijn geëvolueerd om ons als individu in leven te houden en om onze soort – in ons geval de mens – te laten overleven. Een van de allerbelangrijkste vaardigheden hiervoor is het kunnen onderscheiden van biologische beweging en niet-biologische beweging. Als een mensap miljoenen jaren geleden op de savanne het gras afspeurde, wilde hij niet steeds een adrenalinestoot krijgen als de wind het gras deed bewegen, maar wel als het ritselende gras een prooi of een vijand ver raadde. Dezelfde primitieve hersenfunctie zorgt ervoor dat alles wat beweegt zonder een zichtbare externe oorzaak, zoals de wind of zwaartekracht, voor onze hersenen bijzonder fascinerend is. Een robot die beweegt op eigen kracht, die gedreven wordt door artificiële intelligentie, die reageert op zijn omgeving en op mensen, boeit ons mateloos. De robot hoeft niet eens complex te zijn; een stofzuigerrobot slaagt er vaak in om ons uren gekluisterd te laten staren naar hoe hij de woonkamer verkent en zijn weg zoekt rond stoelpoten en tapijten.

Alle dieren kunnen biologische bewegingen van niet-biologische bewegingen onderscheiden, maar wij mensen voegen nog een extra laagje toe wanneer we naar iets kijken wat beweegt. Of we het nu willen of niet, wij interpreteren beweging, en vooral natuurlijk lijkende beweging, steeds als menselijk. We hebben de

neiging om menselijke kenmerken, emoties en intenties te zien in de dingen om ons heen. De wind snijdt, een auto kan lui zijn, en een computer werkt af en toe niet mee. Dat noemen wetenschappers 'antropomorfisme', een samenstelling van de Griekse woorden voor menselijk en vorm, *anthropos* en *morphe*. In 1944 deden psychologen Marianne Simmel en Fritz Heider daarover een experiment: ze maakten een kort filmpje waarin ze zwarte magnetische figuurtjes, cirkels en driehoeken, lieten bewegen op een witte achtergrond. Ze toonden het filmpje aan proefpersonen met de vraag om te beschrijven wat er te zien was. Een makkelijke taak zo bleek. Sommige mensen vonden de cirkel agressief of de driehoek moedig en iedereen zag een verhaaltje waarin de figuurtjes een duidelijke relatie met elkaar hadden. Het merkwaardige hieraan was dat iedereen die het filmpje bekeek, menselijke karaktertrekken zag. En dat is gek, want het gaat enkel om eenvoudige figuurtjes die bewegen op een witte achtergrond. Dit soort antropomorfisme is iets waar robotbouwers nu gretig gebruiken van maken om robots tot leven te wekken.

Pareidolia, wat losjes uit het Grieks vertaald als 'gelijkend op je eigenbeeld', is een andere herseneigenschap waarbij we menselijke trekken zien in dingen die helemaal niet menselijk zijn. Het zien van een gezicht in een wolk, de koplampen en grille van een auto die lijken op ogen en een mond, of een jas aan de kapstok die in het donker op een enge man lijkt. Onze hersenen zijn voortdurend op zoek naar andere mensen en zijn dus hooggevoelig voor alles wat maar enigszins een lichaam of gezicht lijkt te hebben.

Als we een robot bouwen, kunnen we met enkele kleine ingrepen de indruk wekken dat de robot menselijker is dan ze in werkelijkheid is: twee stippen als ogen, een camera op de plaats van de neus, een mondje. Vaak vinden we een robot met een snoet ook aangamer, vriendelijker en intelligenter.

We weten dat we machines vaak als meer menselijk behandelen dan nodig is. Als we gehaast een lift moeten nemen, dan drukken we niet één keer op de knop, maar drukken we een paar keer en net iets harder, omdat we denken dat de lift dan net iets sneller komt. Als een computer niet reageert, gaan we de muis even schudden om de computer wakker te maken. Als een website traag laadt, dan mopperen we tegen de computer. Dit noemen we de *media equation*, het behandelen van computers en robots alsof ze mensen zijn. Dat kan heel gekke gevolgen hebben.

Tijdens een studie aan de universiteit van Stanford in de Verenigde Staten lieten onderzoekers hun proefpersonen een nieuw computerprogramma uitproberen. Om na te gaan wat de mensen dachten van het programma, stelden de onderzoekers een reeks vragen, opnieuw op een computer. Als mensen de vragen beantwoordden op dezelfde computer waarop zij het programma hadden geprobeerd, waren ze beleefder en minder hard dan wanneer ze de vragen op een andere computer beantwoordden. Het leek wel alsof ze de computer niet wilden kwetsen, alsof de computer gevoelens had. Als iemand je vraagt wat je van hun nieuwe kapsel vindt, dan ben je vaak beleefder dan nodig. Hetzelfde gebeurt dus als je mening gevraagd wordt over een computer, en dit effect wordt nog versterkt wanneer we tegenover een robot staan.

De fysieke verschijning van een robot zorgt ervoor dat de robot nog meer aanwezig is dan een computer. Als de robot dan nog eens sociaal overkomt, doordat hij een hoofd heeft of ogen of een kwispelende staart, dan is de robot helemaal niet meer te negeren. Zo'n sociale robot heeft dan niet enkel een fysieke aanwezigheid, maar ook een sociale aanwezigheid en daar kan handig gebruik van gemaakt worden.

Een computer of een robot kan dus heel menselijk en zelfs intelligent lijken door zijn uiterlijk en zijn acties. Maar is dat wel echt zo? Kunnen we artificiële intelligentie (AI) 'intelligent' noemen? Kan een robot een bewustzijn hebben? Moeten we ons zorgen maken over het leervermogen van een robot, want wordt hij daarvoor niet gevaarlijk voor de mensheid? Aan de hand van inzichten uit de wetenschap en de laatste stand van zaken in de robotica probeer ik op deze en andere vragen een antwoord te geven.

ECHTE INTELLIGENTIE EN ARTIFICIËLE INTELLIGENTIE

Wat is intelligentie?

Intelligentie is een moeilijk te vatten begrip. We zijn het erover eens dat de mens intelligent is, maar hoe staat het met andere diersoorten? Een hondenliefhebber vindt vast dat zijn hond intelligent is, net zoals een kattenliefhebber zijn kat een hele Einstein vindt. Maar is een hamster intelligent? Kunnen we een spin intelligent noemen, of waarom niet, een microbe? Het lijkt alsof er een intelligentieschaal is waarop we elke diersoort kunnen plaatsen, met de mens aan het ene uiterste en microben aan het andere uiterste, en alle andere diersoorten daar ergens tussenin. Maar hoe zit het met artificiële intelligentie? Is een computer of een robot intelligent te noemen? Laat ons dit even verkennen met een voorbeeld.

Je gaat zitten met de smartphone in de hand, je opent een app, de cursor knippert uitnodigend en je tikt op het toetsenbord: 'Hoe gaat het?'. Na enkele seconden komt het antwoord: 'Ça va, de zon schijnt en ik doe straks een terrasje.' Het chatgesprek gaat verder en je praat over je plannen voor de dag en over een film die je gisteren gezien hebt, daarna verandert het onderwerp naar muziek en politiek. Je vriend is attent, slim en grappig. Maar stel je nu voor dat je vriend niet van vlees en bloed is, maar dat aan de andere kant van het chatgesprek een computer antwoorde

berekent. Jouw ingetikte karakters komen aan in een serverboerderij duizenden kilometers verderop. Met behulp van informatie over het weer, films, het nieuws, maar ook via jouw voorkeuren, vriendennetwerk en locatiehistoriek berekent een algoritme een gevat antwoord op je input. Zouden we dit algoritme dan intelligent noemen?

Alan Turing, een briljant wiskundige, bouwde tijdens de Tweede Wereldoorlog een machine om Duitse geheimtaal te kraken, die we nu als een van de eerste computers beschouwen. Hij dacht ook na over intelligentie en stelde zich in 1950 de vraag: 'Kunnen machines denken?'. Om die vraag te beantwoorden stelde hij een test voor, nu bekend als de Turingtest. Je krijgt twee computers voor je en op elke computer ga je, enkel gebruik makend van toetsenbord en scherm, een gesprek aan. Aan de andere kant zit ofwel een mens of een andere computer. Als je niet met zekerheid kan zeggen wie de mens en wie de computer is, dan kun je besluiten dat het programma op die computer intelligent is. Turing voorspelde dat een computer eind vorige eeuw een 'gemiddelde ondervrager' tot vijf minuten op het verkeerde been zou kunnen zetten. Een visionaire voorspelling zo bleek, want sinds 1990 bestaat er een jaarlijkse wedstrijd, de Loebnerprijs, waar chatbots worden onderworpen aan een Turingtest. De beste chatbots slagen er inderdaad in om tot enkele minuten te praten over koetjes en kalfjes, en zijn bijna niet te onderscheiden van een mens.

Maar is een gesprek kunnen voeren via een chatsessie wel een teken van intelligentie? Een chatbot ziet, hoort of ruikt niks. Hij kan weinig doen, behalve een reeks karakters op het scherm laten verschijnen. Wij mensen daarentegen kunnen de wereld om ons heen waarnemen met onze zintuigen. We kunnen nadenken, leren en beslissen om iets te doen. En wanneer we iets doen, kunnen we praten, zingen, gezichtsuitdrukkingen vertonen, met onze

handen werken of onze benen gebruiken om ons lichaam naar ergens anders te brengen. We kunnen voorwerpen manipuleren om onze omgeving te veranderen. We kunnen ideeën verspreiden door te praten met anderen of dingen neer te schrijven. Intelligentie is veel meer dan gewoon praten.

Toen computers net bestonden, bijna tachtig jaar geleden, begonnen wetenschappers meteen enthousiast te speculeren over hoe ze die gigantische rekenmachines intelligent konden maken. Ze waren het er allemaal over eens dat ze ambitieus moesten zijn en meteen moesten proberen om computers en software te bouwen die het toppunt van menselijk intellect zouden evenaren.

De uitdaging voor artificiële intelligentie was niet het schrijven van boeken of het oplossen van wiskundige vergelijkingen, maar het spelen van bordspelen. Het bordspel wordt al eeuwenlang beschouwd als een ultieme intellectuele uitdaging. Niet alleen kost het ons tientallen jaren om goed te worden in bordspelen zoals schaak, dammen of go (een Oosters bordspel), een sterk spel spelen vergt creativiteit en inzicht in je tegenspeler. Het is ook mogelijk om te meten hoe goed je bent in het spel. Ofwel win je, ofwel verlies je. En als je meerdere keren speelt, zijn je overwinningen om te zetten in een score. Een computer kan getest worden tegen een mens, maar we kunnen ook twee verschillende programma's tegen elkaar laten spelen om zo te weten te komen welk programma intelligenter is. Wetenschappers dachten ook dat als we eenmaal een computer zouden hebben die uitmuntend schaak speelt, alle andere intellectuele activiteiten, zoals het begrijpen van tekst of camerabeeld, kinderspel zouden zijn.

Zo begon dus een race om AI te bouwen die beter speelde dan de mens. Al in 1959 stelde Arthur Samuel, een computergenie dat toen bij IBM (International Business Machines) werkte, een programma voor dat een voortreffelijk potje dammen speelde. In

Uitgeverij Academia Press
Ampla House
Coupure Rechts 88
9000 Gent
België

www.academiapress.be

Uitgeverij Academia Press maakt deel uit van Lannoo Uitgeverij,
de boeken- en multimediodivisie van Uitgeverij Lannoo nv.

ISBN 978 94 014 6147 4 – D/2019/45/259 – NUR 740

Tony Belpaeme
Artificial intelligence. Op weg naar de artificiële mens?
Gent, Academia Press, 2019, 96p.

Eerste druk, 2019
Vormgeving cover: Studio Lannoo
Vormgeving en zetwerk binnenwerk: Studio Lannoo

© Tony Belpaeme & Uitgeverij Lannoo nv, Tielt

Alle rechten voorbehouden. Niets uit deze uitgave mag worden
verveelvoudigd en/of openbaar gemaakt door middel van druk,
fotokopie, microfilm of op welke andere wijze ook, zonder
voorafgaande schriftelijke toestemming van de uitgever.