

Het onderzoek van Hedderik van Rijn is een zoektocht naar de fundamenteën van het menselijk denken. Hij maakt in zijn modellerwerk veel gebruik van de cognitieve architectuur ACT-R. In deze architectuur spelen symbolische bouwstenen een belangrijke rol. Deze aanpak staat tegenover de connectionistische, gedistribueerde aanpak van neurale netwerken. Karin Zondervan vroeg Hedderik naar zijn successen en tegenslagen en naar zijn standpunt in het symbolisme-connectionisme debat.

INTERVIEW HEDDERIK VAN RIJN DOOR KARIN ZONDERVAN

OP ZOEK NAAR DE FUNDAMENTEN VAN HET DENKEN



Wat zijn de vragen die je probeert te beantwoorden met je onderzoek?

“Ik ben vooral op zoek naar de basisprincipes van menselijke cognitie. Je kunt nooit alles bestuderen en ook niet op alle manieren, maar dat is wel mijn Heilige Graal: hoe werkt de menselijke cognitie nou precies?”

Wat is daar zo spannend aan?

“Menselijke cognitie is iets wat me altijd blijft verbazen. Waarom werkt het op het ene moment zo perfect, terwijl het op een ander moment zo beroerd gaat? De menselijke cognitie is gewoon heel bijzonder in hoe goed en slecht het tegelijkertijd is. Waarom onthoud je soms de meest triviale zaken, terwijl je relevante zaken juist niet kunt onthouden? En kijk naar het alfabet: hoe kan het dat je heel snel weet dat de ‘w’ later in het alfabet zit dan de ‘f’, maar dat het lastiger is om te bedenken welke letter er voor de ‘f’ staat?”

Wat is er “Kunstmatige Intelligentie” aan jouw onderzoek?

“Kunstmatige Intelligentie is een heel breed vakgebied. Aan de ene kant heb je bijvoorbeeld de Pathfinder op Mars. Daarbij maakt het niet zo veel uit of het werkt volgens de menselijke cognitie, als hij het maar doet. Aan de andere kant zit het bestuderen van menselijke cognitie met behulp van technische middelen en principes. Mijn onderzoek bevindt zich aan die kant van het spectrum. Ik ben geïnteresseerd in hoe het zit en niet zozeer in of het werkt.”

Hoe zou je jouw expertise omschrijven?

“Mijn expertise is vooral “verklarend”. Ik maak computationele modellen van gedrag die bedoeld zijn om dat gedrag te verklaren en het liefst te voorspellen. De domeinen waar ik me op begeef, zijn heel breed, veel breder dan je bij sommige andere onderzoekers ziet, die gespecialiseerd zijn op een bepaald onderwerp. Ik ben eigenlijk meer gespecialiseerd op de methode en pas die juist heel breed toe. Ik benader graag allerlei cognitieve taken met een frisse blik. ‘Wat zijn de processen die daar ten grondslag aan liggen?’ En ik probeer ook processen en principes uit het ene domein in het andere domein toe te passen.”

Wat voor praktisch nut kunnen jouw resultaten hebben?

“Cognitieve computermodellen kan je vaak gebruiken om situaties te testen die niet goed testbaar zijn in een menselijke situatie. Als je een cognitieve architectuur hebt die redelijk goed is, dan kan je die cognitieve architectuur een taak laten uitvoeren – als het nodig is zelfs heel

Hedderik van Rijn begon in 1992 als één van de eerste studenten Cognitiewetenschap in Nijmegen. Hij deed een experimenteel afstudeeronderzoek naar het lezen van woorden. Daarna promoveerde hij aan de UvA, waar hij computationele modellen van inductief leren maakte, maar ook van cognitieve ontwikkeling bij kinderen en van het geheugen. Tijdens zijn postdoc project in de Verenigde Staten beperkte hij zich tot één onderwerp: het semantisch geheugen. Maar sinds hij universitair docent cognitief modelleren is geworden aan de RUG zoekt hij weer de breedheid op en maakt hij bijvoorbeeld modellen van hoe mensen tijd inschatten of hoe we het alfabet representeren. Ook de ontwikkelingspsychologie laat hem niet los.

vaak. Zo kan je bestuderen hoe die taak het beste uitgevoerd kan worden, maar ook wat je ervan leert als je heel veel training in de taak krijgt. Natuurlijk zal je altijd ook nog menselijke proefpersonen nodig hebben om je model te valideren, maar als je model eenmaal goed is, dan kan je vrij gemakkelijk dingen veranderen in de taak en kijken hoe dat de prestaties beïnvloedt.”

Dus je gebruikt het computermodel als testpersoon?

“Precies. Met als voordeel dat je in het model kunt kijken om te kunnen zien welke slagen er niet geleerd worden en welke er wel geleerd worden en hoe dat dan precies gebeurt.”

Wat is je droom?

“Eigenlijk is mijn droom dat ik dit altijd leuk blijf vinden. Ik denk niet dat ik het nog meemaak dat de menselijke cognitie uitgeplozen is. Ik denk dat we een heel eind kunnen komen en ik wil ook niet zeggen dat het in principe onmogelijk is, maar ik geloof niet in de voorspelling dat we over vijftig jaar dit punt bereiken zullen hebben. En over vijftig jaar ben ik al ruim met pensioen, dus ik denk dat ik mijn tijd wel vol maak. Ik ben gewoon heel erg geïnteresseerd in menselijke-cognitie principes. Er zijn er nu meer dan ik überhaupt kan onderzoeken. Het uitvinden hoe dingen werken is prachtig en dat wil ik heel graag blijven doen. Er zijn nog heel wat problemen in de ontwikkelingspsychologie, die ik met mijn computationele benadering heel mooi zou kunnen verklaren. Maar hetzelfde geldt voor een aantal geheugentaken en een aantal timing-issues. En voor onderwerpen als impliciet leren. Er zijn zoveel dingen die zo interessant zijn, dat ik gewoon hoop dat het nooit op is.”

Wat was de grootste tegenslag in je wetenschappelijke loopbaan?

“Er is soms wel veel achterdocht en nijd tussen wetenschappers. Sommigen willen niks vertellen over hun werk, want wie weet ga jij er wel met hun experimenten van door. Ik had me de onderzoeksweld veel meer voorgesteld als een soort paradijs: dat je met zijn allen slim bent en dat je met zijn allen probeert de kennis die we hebben verder te brengen. Maar die paradijselijke omgeving blijkt ook dorre stukken te hebben. Er zijn genoeg onderzoekers die alleen in hun eigen theorie geloven en die helemaal niet open staan voor alternatieve ideeën.”

Dat had je wel meer verwacht in de wetenschappelijke wereld?

“Ja. In feite streef je allemaal hetzelfde doel na. Je hebt een theorie over een fenomeen en je laat zien dat de theorie van iemand anders over dat fenomeen niet klopt. Je hebt het er met die persoon over gehad en hij is het ermee eens dat jouw voorspellingen beter kloppen dan die van hem. Maar toch hoor je zo iemand even later in een lezing beweren dat zijn verklaring de beste is. Dat is wel logisch, maar het is ook jammer. De wetenschap is menselijker dan je soms zou willen.”

Waarin verschil jij van andere cognitiewetenschappers?

“Samen met mijn directe collega's ben ik vooral bezig op een symbolische manier te modelleren. Het leuke van die benadering is dat je echt weet wat er gebeurt. Als je het model openmaakt, dan kan je heel duidelijk zien wat de bijdrages van de verschillende modules zijn aan het antwoord.”

Jij werkt dus vanuit een symbolisch uitgangspunt. Speelt het debat tussen symbolisme en connectionisme een rol in jouw onderzoek?

“Naar mijn idee is het is net als de discussie over of je licht moet zien als deeltjes of als golven: het klopt beide niet, of beide wel, afhankelijk van hoe je ernaar kijkt. Een argument dat symbolisten vaak horen van connectionisten is dat hun modellen niet biologisch plausibel zijn. Dat we in ons hoofd geen productieregels hebben. Dat klopt. Maar neurale netwerken zijn ook niet per definitie cognitief plausibel. Voor ieder probleem worden eigen input- en outputunits bedacht en de netwerkstructuren zijn ook steeds weer anders. Ook gebruiken netwerkbouwers vaak voorgestructureerde input. Het wordt vaak gepresenteerd als een kwalitatief onderscheid: neurale netwerken zijn cognitief plausibel en symbolische modellen zijn dat niet. Maar volgens mij is het verschil vooral kwantitatief van aard. Je zou kunnen zeggen dat bepaalde typen neurale netwerkmodellen, met name die modellen die ook specifiek letten op de biologische plausibiliteit, inderdaad meer op echte hersenen lijken. Maar het meest plausibele model van het menselijk brein is een implementatie van het gehele brein. Dat is zo complex, dat we niet precies weten hoe het werkt. We weten dankzij EEG en MRI veel over wanneer en waar iets gebeurt, maar hoe dat nou precies in elkaar zit, dat weten we niet. Je moet dus wel versimpelen en abstraheren. En dan is een abstractie naar het niveau van de huidige typen neurale netwerken in feite net zo erg als een abstractie naar het niveau van bijvoorbeeld ACT-R of bepaalde typen mathematische

“Het debat tussen symbolisme en connectionisme is net als de discussie over of je licht moet zien als deeltjes of als golven: het klopt beide niet, of beide wel, afhankelijk van hoe je ernaar kijkt.”

modellen.”

Kunnen we cognitie tegelijkertijd zien vanuit het deeltjesperspectief, maar ook vanuit het golfperspectief?

“Het hangt ervan af welke taak je wilt bestuderen. Ik denk dat voor veel van de hogere cognitieve acties op dit moment de neurale-netwerkenmethode minder passend is. Dan moet je vaak alsnog zoveel kunstgrepen uithalen dat het niet meer duidelijk is wat je model precies doet en voorspelt. Of er moeten zoveel cognitief implausibele aspecten toegevoegd worden, dat het – in elk geval soms – minder plausibel wordt dan een symbolisch model.”

En dan ben je dat voordeel dus alweer kwijt?

“Inderdaad. Daartegenover staat dat je met een symbolisch model gewoon kan zien wat er gebeurt. Maar je moet je wel afvragen: wat mag je doen in één stap? En is het wel plausibel dat dat in één stap gebeurt? Een symbolisch model mag je bijvoorbeeld nooit laten werken met een soort lijststructuur als geheugen, want mensen hebben geen perfecte lijststructuur. Mensen moeten moeite doen als ze een boodschappenlijstje willen onthouden.”

Maar er zijn dus wel symbolische modellen gebouwd die met lijst-gebaseerde geheugens werken?

“Ja, lang geleden. Maar dat is gewoon niet correct en dat weten we ondertussen ook wel. Het voordeel van zo’n misstap is overigens wel dat het inzicht geeft in hoe mensen het dan wel doen: ze groeperen bijvoorbeeld concepten om alsnog die lijst te onthouden. Bijvoorbeeld: ik weet dat ik vanavond cake ga bakken, dus dan heb ik eieren, boter, suiker en bloem nodig. Je moet je bij de symbolische aanpak simpelweg afvragen: zou dit op een realistische manier in een hersencorrecte neurale netwerkvariant geïmplementeerd kunnen worden? En dat kan je doen zonder dat je het daadwerkelijk zo gaat implementeren. Want soms weet je gewoon: zo’n hersencorrecte implementatie zou waarschijnlijk wel lukken, als ik daar de rest van mijn carrière de tijd voor neem...”

En dan besteed jij dus liever je tijd aan het modelleren van cognitieve taken op een wat abstracter niveau?

“Als je kijkt wat er nu geïmplementeerd is op echt biologisch

plausibel niveau – waarbij je dus echt kan zeggen: dit onderdeelje van het model is gelijkwaardig aan dat neuron of dat clusterje van neuronen in het brein – dan gaat het meer over zaken als: kan iemand herkennen dat er een streep staat? Dat is nog lang niet in de buurt van bijvoorbeeld het herkennen van woorden. Mijn idee is dat je principes moet nemen uit hoe wij weten dat de menselijke hersenen werken en dat je die principes moet gebruiken om een model te maken waarmee je verder kan.”

Waarom zijn neurale netwerken niet per definitie biologisch plausibeler? We weten toch dat onze hersenen parallel en gedistribueerd werken?

“Mensen denken vaak: het is een neurale netwerk, dus het is plausibel. Maar dat is gewoon onzin. Op het internet kun je bijvoorbeeld het spel 20-questions¹ spelen tegen een neurale netwerk. Ze verkopen het alsof de computertegenstander het spel “op een menselijke manier” speelt, maar het blijkt gewoon een soort boomstructuur te zijn die in een neurale netwerk gegoten is. Het is een correlatienetwerk met knopen die keuzes en woorden representeren. Maar als je een knoop mag hebben dat een woord representeert, dan ben je hetzelfde niveau bezig als symbolische modellen, waarbij één symbool een woord representeert.”

Waar gaat het debat tussen connectionisten en symbolisten heen?

“Ik denk dat we vanuit drie kanten naar elkaar toe moeten werken. Vanuit de symbolische beschrijvingen op het complexere gedragsniveau, vanuit het lage gedistribueerde niveau van neurale netwerken en vanuit het onderzoek dat metingen doet aan echte hersenen. Zodra die drie naar elkaar toe gaan, krijg je pas echt een soort Heilige Graal. Op dat moment moet ook de Nobelprijs voor de Cognitiewetenschap worden ingesteld. De persoon die voor het eerst deze drie startpunten goed samen kan brengen verdient die prijs.”

Dus het moet geen debat zijn, maar meer een situatie waarin verschillende onderzoekers vanuit verschillende startposities naar hetzelfde doel werken?

“Heel veel mensen gedragen zich echter in het openbaar alsof ze oogkleppen op hebben. Ze kijken alleen in de richting van hun



HEDDERIK VAN RIJN

eigen Heilige Graal. En ze zetten zich daarbij af tegen de andere groepen, die andere uitgangspunten en methodes hebben. Want die zijn nou eenmaal niet juist bezig.”

Die praktijk druist in tegen jouw paradijselijke idee van wetenschappers die met verschillende benaderingen en verschillende technieken samenwerken aan hetzelfde bouwwerk?

“Ja. Ik probeer daar ook echt wat aan te doen. Eén van de dingen waar ik op dit moment mee bezig ben, is een idee uit de neurale-netwerken literatuur toepassen in de symbolische architectuur ACT-R en kijken of dat tot een verbetering leidt.”

Hoe zie je de toekomst van het cognitief modelleren?

“Zeer positief. Het is duidelijk geworden dat het maken van computermodellen van menselijk cognitief gedrag een zeer nuttige exercitie is, als je wilt weten of jouw verklaring van een bepaald fenomeen de juiste is. Als je alleen een beschrijvend model maakt, dan is het heel moeilijk om zeker te weten dat de causale

structuren die je beschrijft ook daadwerkelijk de juiste zijn. Als je een computermodel maakt en het klopt niet, dan merk je meteen dat je model niet het juiste gedrag vertoont. Je wordt met je neus op de feiten gedrukt. Cognitief modelleren is wat dat betreft ook wel een beetje een pokkebaan. Vaak genoeg moet je concluderen: dit is het dus ook niet helemaal. Mijn implementatie van de beschrijvende theorie klopt niet. Maar uiteindelijk kun je veel specifieker zijn in je voorspellingen van menselijk gedrag. En je weet veel beter weten waar je de interessante vervollexperimenten moet zoeken. Computermodellen dwingen je ook om alles heel expliciet maken. Je kunt je niet achter vage beschrijvingen verschuilen. Als je een cognitief model maakt, moet je echt met de billen bloot.”¹

1. zie <http://www.20q.nl>

OP ZOEK NAAR DE PERFECTE MASTER?

(Bijna) klaar met je bachelor, op zoek naar de perfecte master-opleiding? Wij bieden je een overzicht van alle Nederlandse AI en AI-gerelateerde masters. Check nu onze nieuwe mastersite op:

<<http://www.connectie.org/master.php>>