

De keuze van Johan van Benthem

Eva van den Broek

In ieder nummer van *De Academische Boekengids* vertelt een wetenschapper over de boeken die een inspiratie vormden in leven, opleiding of onderzoek.

Johan van Benthem (1949), universiteitshoogleraar aan de Universiteit van Amsterdam en hoogleraar filosofie aan Stanford University, studeerde wiskunde en filosofie. Nu werkt hij op het grensvlak van logica, informatica en speltheorie. Zelf omschrijft hij zijn vakgebied als 'cognitiewetenschap zonder feiten'.

Deze vijf boeken omspannen de dimensies van mijn intellectuele ruimte. Was het niet de Nederlandse godgeleerde Geert Groote die zei dat hij in het leven van alles geprobeerd had, maar dat hij het ware geluk alleen kon vinden met een boekje in een hoekje, *in angello cum libello*? In die uitspraak kan ik me goed vinden.

A History of Greece van John Bury en Russell Meiggs beslaat ongeveer de zesde tot de derde eeuw voor Christus. Het is een van de eerste boeken die mijn zelfbeeld voorgoed veranderd hebben. Ik was veertien toen ik het op de middelbare school als prijs kreeg. Ineens was er het besef dat ik leefde in een tweeduizend jaar oude gemeenschap van gelijkgestemde mensen. Dat gevoel heb ik in mijn werk nog steeds: dat ik een traditie voortzet die veel groter is dan ikzelf.

Wat me in het boek vooral trof, was de moderniteit van de teksten, de overeenkomsten tussen het wereldbeeld van de schrijvers en dat van mij. Zo zegt Thucydides, een van Bury's helden, over een partijtivist in Corcyra: 'De oligarchen haatten de democraten, de democraten haatten de oligarchen. Maar het meest van al werden diegenen gehaat die geen partij kozen.' Dat soort inzichten maakte indruk op me. Met mensen die zo denken en schrijven voelde ik me verwant.

Wat mij daarnaast aantrok in de Griekse cultuur was het contrast tussen die denkwereld vol *sophistication*, intellectuele diepgang en humor en de Bijbel. Ik kom uit een calvinistisch gezin. Als kind vroeg ik me al af waarom we maar één god hebben. Volgens de Bijbel is alleen het monotheïsme zaligmakend, maar het leek me juist een kwestie van intellectuele souplesse en tact tegenover andersdenkenden om een aantal verschillende goden te accepteren.

Het is een cultuurramp van de eerste orde geweest dat van de Joodse literatuur alleen de Bijbel het heeft gered. Nu citeren we als 'bron van onze cultuur' een fundamentalistisch partijprogramma. Juist de analogieën en mengvormen tussen godsdiensten kunnen inzichtelijk zijn – een houding die ik ook in mijn werk aanneem als tegenwicht tegen de veelvuldige kerkvorming binnen de wetenschap.

'LANGE TIJD HEEFT DE NATUURLIJKE TAAL EEN SLECHTE PERS GEKREGEN.'

Een boek dat mij tot de logica heeft bekeerd, was *Gödel's Proof* van de wetenschapsfilosoof Ernest Nagel en wiskundige James Newman. Tijdens mijn studie kreeg ik het idee dat logica iets bijzonders was. Ik hoorde van het bestaan van Gödels onvolledigheidsstelling, die iets heel fundamenteels beweerde, namelijk dat er grenzen zijn aan wat valt te bewijzen in wiskundige taal. En blijkbaar kon je die grenzen uitdrukken – en zelfs bewijzen – in dezelfde taal. Dat was vreemd. Als je buiten een systeem staat, kun je er iets over zeggen, maar toch niet als je er binnenin staat?

Deze ervaring werd nog versterkt door de confrontatie met de zogeheten leugenaarsparadox. Die luidt: 'Deze zin is onwaar.' Nu zijn er twee mogelijkheden. Stel, de bewering is waar: datgene wat de zin zegt, is dan onwaar, dus spreekt hij zichzelf tegen. Stel, de bewering is onwaar, dan heeft de zin dus gelijk – want dat zegt hij nou juist! Maar dan is hij dus tóch waar. Met andere woorden: de zin is slechts waar als hij niet waar is. Dit is een heel simpele en directe paradox. Het heeft iets schokkends dat hij met zulke eenvoudige ingrediënten ontstaat: alleen de begrippen 'waar' en 'zin', met misschien een beetje logica over als-dan-redeneringen. Er lijkt dus een probleem te zitten in het fundament van ons denken. Voor mij was deze paradox een soort vulkaan, een klein puntje op het aardoppervlak, maar daaronder vermoed je dat het kolkt.

Gödels gezichtspunt stelt hem vervolgens in staat deze leugenaarsparadox om te zetten in een valide wiskundig inzicht. Hij vervangt het begrip 'waar' door 'bewijsbaar'. Gödels zin zegt niet 'ik ben onwaar' maar 'ik ben onbewijsbaar'. Gödel laat zien dat de wiskunde een tegenspraak bevat als die zin bewezen kan worden. Maar als dat niet lukt, is hij kennelijk waar – en dan is de wiskunde onvolledig, want sommige ware uitspraken zijn dan niet te bewijzen. Met andere woorden: wiskundige theorieën zijn of inconsistent of onvolledig. Zulke finesses hebben op mij een blijvende indruk gemaakt.

'DE LEUGENAARSPARADOX IS EEN VULKAANKRATER: JE VERMOEDT DAT HET ERONDER AAN HET KOLKEN IS.'

Gödel deed met zijn wiskundige benadering een stap weg van de natuurlijke taal. Lange tijd heeft die een

slechte pers gekregen. De filosofen Ludwig Wittgenstein en Bertrand Russell, bijvoorbeeld, vonden onze dagelijkse taal maar onsystematisch en misleidend als weergave van betekenisvolle uitspraken. Zelfs de taalkundige Noam Chomsky liet zich niet echt in met betekenis: hij benaderde taal syntactisch en keek vooral naar de vormelijke eigenschappen ervan. De Amerikaanse logicus Richard Montague liet als eerste zien dat natuurlijke taal, inclusief de semantiek, de betekenisstructuur, heel goed te behandelen valt met wiskundige methoden.

In de jaren tachtig ontstond aan de Universiteit van Amsterdam een bolwerk van mensen die de dagelijkse taal bestudeerden op basis van het onderzoek van Montague. Naast het beschrijven van de specifieke betekenis van woorden en constructies stelden zij al snel ook algemenere vragen. Bijvoorbeeld: wat zijn de belangrijkste structuren die elke taal moet kunnen weergeven? Als logicus denk ik daarbij allereerst aan de voegwoorden 'en' en 'of', en aan hoeveelheidsuitdrukkingen zoals 'alle', 'sommige' en 'de meeste'. Je kunt bijvoorbeeld onderzoeken hoeveel woorden verschillende talen hebben voor het uitdrukken van 'alle' of 'meeste'. Maar talen kunnen dat op heel verschillende manieren aangeven. Dus moet je op een hoger abstractieniveau dan de vorm kijken. De logisch-wiskundige methode van Montague gaf dat vereiste hogere abstractieniveau.

Het werk van Montague riep nog een tweede vraag voor me op: waarom zijn logici niet overgestapt op redeneren en communiceren in een formele taal? Zelfs zij – degenen die er dagelijks mee werken en er de gereedschappen en training voor hebben – maken hiertoe geen aanstalten. Misschien komt dat doordat eenvoudig redeneren snel gaat in natuurlijke taal, zodat er een groot gebied is waarin het niet nodig is om over te gaan tot een formele taal.

'Natuurlijke' gevolgtrekkingen kun je bijvoorbeeld expliciet maken met eenvoudige wiskundige modellen. Als een zin waar is, kun je een predicaat, oftewel een bepaling, soms vervangen door een andere bepaling met een ruimere of engere betekenis. Uit de zin 'Alle aanwezigen liepen weg' kun je bijvoorbeeld opmaken dat ook alle *jongens* wegliepen. Andersom, veralgemeniserend, concludeer je dat alle aanwezigen ook *bewogen*. Dit soort gevolgtrekkingen met predicaatvervanging gebruiken we voortdurend. Bij 'sommige' kan bijvoorbeeld zowel het werkwoord als het onderwerp worden verruimd. 'Sommige jongens dansen': dus sommige *mensen* dansen, en sommige jongens *bewegen*.

Blijkbaar zitten in de grammatica allerlei redeneermethoden verborgen die wij onbewust en heel snel gebruiken. In de jaren tachtig heb ik veel gewerkt aan 'natuurlijke logica', snelle redeneervormen die dicht bij de natuurlijke taal staan. Nu krijgt dat gebied weer aandacht, ook binnen de informatica en cognitiewetenschap.

De volgende sprong in mijn onderzoek is sterk beïnvloed door *Knowledge in Flux* van Peter Gärdenfors, een Zweedse filosoof en logicus, en tegenwoordig invloedrijk cognitiewetenschapper. Zijn spectaculaire inzicht was dat allerlei informatieverwerkende processen die traditioneel niet tot de logica behoren wel degelijk tot object van logisch onderzoek kunnen worden gemaakt. Het gaat bijvoorbeeld niet alleen om bewijzen en consistent kennis op kennis stapelen. Juist onze reacties op informatie die strijdig is met wat we tot nu toe geloofden, vormen een wezenlijk kenmerk van intelligent gedrag.

Gärdenfors gaf een formele theorie van *belief revision* – een proces waarvan je zou hopen dat het zich wat vaker voordeed bij sommige mensen. Hoe dat revisieproces werkt, hoe je je systeem van overtuigingen weer consistent maakt, daarvoor laat Gärdenfors' theorie de ruimte: van radicale geloofsherziening tot conservatieve minimale aanpassing.

Sinds ik Gärdenfors heb gelezen, zie ik de logica als een algemene theorie van informatieverwerkende processen, van gevolgtrekkingen tot waarneming, en van consistent kennis vergaren tot spectaculaire episodes van geloofsherziening. Gärdenfors liet concreet zien dat het mogelijk is om met logische technieken niet alleen de statische informatie-inhoud in formules te beschrijven maar soms ook het hele dynamische proces daarachter. Van dit dubbele perspectief zie je overigens de basis al in de natuurlijke taal. Neem een woord als 'bewijzen': het verwijst zowel naar de dynamische activiteit als naar de statische neerslag daarvan, de stukken tekst.

Deze visie op logica uitleggen is zelf een dynamisch proces. In het Amsterdamse science center NEMO heb ik ooit een *Kids' lecture* over logica gegeven. Logica is moeilijk uit te leggen, zeker aan kinderen, maar ik varieerde op een voorbeeld dat mijn geliefde Griekse voorouders tweeduizend jaar geleden al gaven. Een hond volgt een spoor met zijn neus. Het pad splitst zich, de hond snuffelt links en loopt dan naar rechts. Waarom doet hij dat, waarom snuffelt hij rechts niet? Omdat hij weet dat het spoor doorloopt en als het niet naar links gaat, moet het wel naar rechts gaan. Je ziet hier een klassieke logische gevolgtrekking: 'A of B, niet A, dus B.' En tegelijkertijd ook een heel moderne, want dit is natuurlijk het sudokupatroon waar half Nederland mee puzzelt.

Observaties veranderen stapsgewijs de informatietoestand van personen; soms op een eenvoudige, consistente manier, soms schoksgewijs. Gärdenfors merkte overigens op dat ook de wetenschappelijke activiteit zelf onder deze analyse valt. Het vinden van een contradictie is geen ramp, want procesmatig altijd interessant. Een essentieel onderdeel van wetenschappelijke rationaliteit is niet de correctheid van onze theorieën, maar de kwaliteit van onze herziening als ze incorrect blijken.

'GELOOF-SHERZIENING ZOU ZICH BIJ SOMMIGE MENSEN WAT VAKER MOGEN VOORDOEN.'

Nadat Gärdenfors het werkterrein van de logica vergroot had door zijn dynamische wending, markeerde *A Course in Game Theory* van de economen Martin Osborne en Ariel Rubinstein voor mij de stap naar het beschrijven van de interactie tussen verscheidene personen. Redeneren en observeren kun je nog beschouwen als cognitieve activiteiten van één persoon, maar zowel natuurlijke taal als logica gaat natuurlijk evenzeer, en misschien nog wel meer, over het stellen van vragen en het argumenteren met anderen. Deze sociale functies van de logica waren in wezen al in de oudheid aanwezig. In essentie berusten dynamische, logische processen op *interactie*: we moeten 'many mind problems' oplossen, net zoals de natuurkunde doet voor 'many body problems'. Soms ontstaat deze interactie zelfs in één hoofd. Op die manier studeerde ik vroeger: wandelend over straat in Amsterdam vertelde ik aan mezelf wat ik wel en niet begreep van een college, soms tot bezorgdheid van voorbijgangers.

Maar hoe kun je die interactie nu precies in logische termen beschrijven? Daarvoor moet de logica nieuwe allianties aangaan, bijvoorbeeld met de speltheorie. Osborne en Rubinstein schreven een vrij compromisloze wiskundige introductie tot de speltheorie, met aardige dialogen tussen de auteurs als ze het onderling oneens zijn. Het is geen populair boek, maar mij toch op het lijf geschreven. Ik heb eruit geleerd dat de speltheorie, die belangrijke inzichten geeft in het nemen van beslissingen door mensen, de logica kan verrijken door de doelen en voorkeuren van de 'spelers' wiskundig te beschrijven.

In een conversatie heeft de overdracht van informatie bijvoorbeeld altijd een bepaald doel. Als jij mij een vraag stelt, zit daar voor mij vaak een waaromvraag in verborgen: waarom wil je dat weten? Met de vraag naar de doelen van de verschillende spelers zit je dicht bij de speltheorie - die draait om optimale interactieve strategieën tussen spelers, rekening houdend met het doel van de ander. Je kunt spelmodellen maken voor vele logische begrippen, waaronder argumentatie, om de onderliggende interactie te begrijpen. Wij doen aan het Institute for Logic, Language and Computation tegenwoordig allerlei onderzoek dat de speltheorie koppelt aan logica en taalkunde.

Dat is niet alleen een kwestie van iets leren van de speltheorie, er blijkt ook ruimte voor verkeer in de omgekeerde richting. De speltheorie draait om het vinden van 'evenwichten' waarin iedereen, gegeven de situatie, datgene doet wat voor hem of haar het beste is. In zo'n evenwicht kan niemand zijn situatie verbeteren door als enige af te wijken van de strategie. Het lijkt echter alsof er in de speltheorie een gat zit tussen het wiskundige apparaat dat die evenwichten berekent en ons intuïtieve begrip daarvan, hoewel je de Nobelprijs Economie kunt winnen zonder dit uit te leggen (in 2005 ging de Nobelprijs voor economie naar speltheoretici, EvdB). In dat gat is ruimte voor de logica. De speltheorie heeft geen goed model van de spelers; de redeneerstappen en andere dynamische processen ontbreken daarin. Hier kunnen logica en speltheorie gezamenlijk voortgang boeken.

'VOOR VEEL COLLEGA'S BEN IK DIVERSE BRUGGEN TE VER UIT MIJN VLIEGTUIG GESPRONGEN.'

Met deze opvattingen ben ik in de logica een nieuwlichter die voor veel collega's diverse bruggen te ver uit zijn vliegtuig is gesprongen. Maar ik zie deze nieuwlichterij ook als een terugkeer naar de bronnen. De logische traditie die teruggaat tot de oudheid kwam deels voort uit wiskundige bewijzen zoals die van Euclides. Daarnaast denk ik dat Plato's dialogen een belangrijke invloed hebben gehad. Waarschijnlijk is de logica ontstaan uit het juridische en politieke debat, de vorm van argumentatie en de patronen in interactie. Negatie, in de logica het ontkennen van een standpunt, is in deze opvatting ineens een centraal cognitief vermogen: een verandering van perspectief.

Dat ik mij ten slotte nog steeds beroep op de oudheid en de continuïteit met mijn voorgangers van tweeduizend jaar geleden: daar sluit voor mij de cirkel van mijn vijf boeken.

Prof. dr. J.F.A.K. van Benthem is universiteitshoogleraar aan de Universiteit van Amsterdam op het gebied van de logica en haar toepassingen. Tevens is hij houder van de Bonsall Visiting Chair of Humanities aan Stanford University (Californië) en grondlegger en directeur van het Institute for Logic, Language and Computation (ILLC). In 1996 ontving hij de NWO-Spinozapremie voor zijn gehele wetenschappelijke onderzoek. In 2008 verschijnt het *Handbook of Philosophy of Informatics* (Elsevier, met mederedacteur P. Adriaans).

De keuze van:

A COURSE IN GAME THEORY
door **Martin J. Osborne en Ariel Rubinstein**
MIT Press. Cambridge 1994., € 41,80

A HISTORY OF GREECE TO THE DEATH OF ALEXANDER THE GREAT
door **John B. Bury en Russell Meiggs**
Macmillan 1977 (Oorspr. 1900), € 41,75

FORMAL PHILOSOPHY - SELECTED PAPERS OF RICHARD MONTAGUE
door **Richard Montague (Richmond H. Thomason red.)**
Yale University Press. New Haven 1974.

GÖDEL'S PROOF
door **Ernest Nagel en James R. Newman**
New York University Press. New York 2001.
(Oorspr. 1958), € 20,30

KNOWLEDGE IN FLUX - MODELING THE DYNAMICS OF EPISTEMIC STATES
door **Peter Gärdenfors**
MIT Press. Cambridge 1990. (Oorspr. 1988)