

FILOSOFIE VAN DE KUNSTMATIGE INTELLIGENTIE

De klassieke filosofie van de kunstmatige intelligentie of ook filosofie van de AI¹ werd ooit door Feigenbaum en Feldman als aparte deeldiscipline op de kaart gezet toen zij in de systematische bibliografie van hun verzamelbundel *Computers and Thought* een rubriek *L* met vier kennistheoretische vragen (*Epistemological Questions*) opnamen:²

L₁ What Can a Machine Know?

L₂ Can Machines Think? Nature of Intelligence

L₃ Free Will in Man and Machines

L₄ The Mind-Brain Problem

Deze vragen zijn misschien enigszins misleidend. We zeggen bijvoorbeeld van een schaakcomputer al dat-ie ‘nog aan het denken is’ zolang-ie aan zet is en nog geen zet heeft gedaan, dus waarom zouden machines niet kunnen denken? Waarom zouden ze anders kunnen rekenen en zelfs schaken? Niet dat de kwestie van het gebruik van antropomorfe termen bij machines niet als problematisch kan worden ervaren, maar het gaat bij filosofie van de AI om typisch ‘filosofische’ vragen die als volgt kunnen worden herformuleerd:

L1* Zijn er grenzen aan ‘machinale kennis’ die er niet zijn voor menselijke kennis?

L2* Zijn er grenzen aan ‘machinale procedures’ die er niet zijn voor menselijke intellectuele activiteiten?

L3* Zijn er grenzen aan ‘machinale keuzeregels’ die er niet zijn voor menselijke beslissingen?

L4* Zijn er grenzen aan de mogelijkheden van machines die er niet zijn voor menselijke hersenen?

¹ ‘Klassieke’ filosofie van de AI wordt gedefinieerd door de navolgende vragen L1 t/m L4. Voor een meer ‘moderne’ problematiek verwijs ik naar Klaus Mainzer, *KI – Künstliche Intelligenz. Grundlagen intelligenter Systeme*. Darmstadt: Wissenschaftliche Buchgesellschaft, 2003.

² Edward A. Feigenbaum and Julian Feldman (eds.), *Computers and Thought*. New York etc.: McGraw-Hill, 1963, pp.471-472. De auteur van de bibliografie was Marvin Minsky.

L1*, L2*, L3* en L4* hebben elk twee kanten: aan de ene kant gaat het om toeschrijvingen aan machines, aan de andere kant aan toeschrijvingen aan mensen.

De geschiedenis laat zien dat filosofische ingestelde personen al over zulke problemen hebben nagedacht, ondanks het feit dat zij hetzij alleen rekenmachines kenden (Pascal, Leibniz, Peirce), of digitale computers nog in de kinderschoenen stonden (Turing, Bunge, A. D. de Groot), kennelijk omdat zij zo'n goed inzicht meenden te hebben in de *mogelijkheden die in principe* door zulke apparatuur werden gecreëerd. En gaat filosofie – vrij naar Wittgenstein³ – niet over elke mogelijkheid en zijn niet alle mogelijkheden haar 'feiten'?

Vandaag de dag zijn computers – en ten dele ook robots – zover ontwikkeld dat de claims uit de begintijd beter lijken te kunnen worden beoordeeld. Het gaat nu niet alléén meer om mogelijkheden, maar ook om onfilosofische feitelijkheden. Het laatste geldt overigens zowel voor artefacten als voor mensen voorzover de laatsten bepaalde dingen in feite niet meer 'kunnen', om de eenvoudige reden dat zij zoiets niet meer hoeven te leren omdat computers en robots dat veel beter doen.⁴ Aan de andere kant kunnen mensen heden ten dage veel meer dan vroeger, niet alleen omdat zij de hulp van computers en robots kunnen inroepen, maar ook omdat de computerwetenschappen hun nieuwe oplossingsstrategieën en -methoden aan de hand hebben gedaan.

Aldus gaat filosofie van de kunstmatige intelligentie ook over 'onzelf', om het quasi-filosofisch uit te drukken. Vandaar dat in de diverse filosofische opvattingen over kunstmatige intelligentie bepaalde mensbeelden tot uitdrukking komen. Dat was trouwens al op te maken uit de bovengenoemde herformuleringen van de vragen van Feigenbaum en Feldman. Welnu, om de heersende tegenstellingen duidelijk te laten uitkomen, onderscheid ik sinds jaar en dag twee 'ideaaltypen', neo-intuitionisme en neo-mechanicisme.⁵ Argumenten van enkele van hun vertegenwoordigers zullen vervolgens de revue passeren.

Ik ben begonnen de vragen van Feigenbaum en Feldman te herformuleren. Maar ook de herformuleerde vragen behoeven nog toelichting. Wat wordt er bedoeld met "machinale kennis", wat met 'machinale procedures' en wat met 'machinale keuzeregels'? Antwoorden hierop kunnen mijns inziens het beste gegeven worden in termen van probleemoplossingen. Dit is trouwens mijn algemene strategie. Je hoort wel eens filosofen vragen of computers wel een 'bewustzijn' kunnen hebben. Los van de kwestie wat er onder 'bewustzijn' moet worden verstaan, kan de vraag onmiddellijk worden teruggespeeld: geef eerst een

³ Ludwig Wittgenstein, *Logisch-Philosophische Abhandlung. Annalen der Naturphilosophie* 14, 1921, p. 199 (2.0121).

⁴ Vgl. Henk Visser, *Wat mensen niet meer kunnen doen. BNAIS-lezing 2003.*

⁵ Vgl. Henk Visser, *Humanistic views on man and computer. Rekenschap*, 39, 1992, pp. 97-103.

probleem dat met behulp van bewustzijn kan worden opgelost, laat vervolgens zien dat dit ook niet zonder bewustzijn kan, en toon tenslotte aan dat computers het probleem wel of niet kunnen oplossen. Het zal duidelijk zijn dat de tweede stap het moeilijkst is en dit brengt mij meteen bij een cruciaal verschil van opvatting onder filosofen van de AI: hetzij benadrukken zij de *processen* waarmee mensen en computers problemen oplossen, of zij concentreren zich op de probleemoplossingen in de zin van de *resultaten*. En in het algemeen zijn proces en resultaat niet equivalent, hoewel het ook mogelijk is resultaten onlosmakelijk verbonden te achten met de processen waardoor zij tot stand zijn gekomen. Dat werd ooit door Wittgenstein gedaan, bij zijn beoordeling van problemen in de filosofie van de wiskunde.⁶ Daar is in dat geval veel voor te zeggen. Ik herinner mij in dit verband een uitspraak van de wiskundige Heyting, die filosofische standpunten alleen het bespreken waard vond als de manier van redeneren er door beïnvloed wordt.

Over filosofie van de wiskunde gesproken, er is een opmerkelijke parallel tussen opvattingen daarin en filosofische posities over computers. Maar daarover aanstonds.

Allereerst wil ik wijzen op het gemaakte onderscheid tussen ‘kennis’ en ‘procedures’. Dit houdt in mijn terminologie in dat problemen opgelost kunnen worden met behulp van ‘principes’, bijvoorbeeld bij wiskundige problemen in de vorm van al bewezen stellingen, en, inderdaad, ‘procedures’, bijvoorbeeld redeneerwijzen zoals bewijzen uit het ongerijmde en *ex hypothesi*, maar ook strategieën zoals ‘terugwerken’ in plaats van vanuit de gegevens. Dit kan met name plaats vinden bij zgn. toestandruimteproblemen door te beginnen met een eindsituatie in plaats van met de begintoestand. U heeft dit in uw kindertijd vast wel eens gedaan bij puzzeltjes zoals ‘wie van de drie heeft beet’.

Een mooi voorbeeld van zulk terugwerken levert een opgave van Euler,⁷ die ik u niet wil onthouden:

Drie mensen spelen om geld. Na het eerste spel verliest de eerste speler aan de andere spelers net zoveel geld als zij op dat moment bezitten, na het tweede spel ondergaat de tweede speler dit lot, en na het derde spel de derde. Dan blijken alle drie het zelfde bedrag te bezitten, namelijk 8 florijnen. Hoeveel had eenieder aan het begin?

⁶ Vgl. Ludwig Wittgenstein, *Bemerkungen über die Grundlagen der Mathematik (Remarks on the Foundations of Mathematics)*. Edited by G. H. von Wright, R. Rhees and G. E. M. Anscombe. Oxford: Basil Blackwell, 1956, p. 26. Vgl. Ludwig Wittgenstein, Logisch-Philosophische Abhandlung, *Annalen der Naturphilosophie* 14, 1921, p. 254 (6.1261).

⁷ Vgl. Leonhard Euler, *Vollständige Anleitung zur Algebra*. Neue Ausgabe. Leipzig: Philipp Reclam jun., z.j., pp. 241-243.

Om dit probleem op te lossen, kunnen mensen – die deze beschrijving begrijpen en de nodige wiskundeopleiding hebben gehad en het volgende nog kunnen – beginnen met de bezittingen in euros aan het begin op x , y , en z te stellen, om vervolgens algebraïsch de resultaten van de opeenvolgende spelen te bepalen:

$$\begin{array}{c}
 x, y, z \\
 x - y - z, 2y, 2z \\
 2(x - y - z), 2y - (x - y - z) - 2z, 4z \\
 4(x - y - z), 2(2y - (x - y - z) - 2z), 4z - 2(x - y - z) - (2y - (x - y - z) - 2z)
 \end{array}$$

en tenslotte het volgende stelsel van vergelijkingen trachten op te lossen:

$$\begin{array}{l}
 4(x - y - z) = 8 \\
 2(2y - (x - y - z) - 2z) = 8 \\
 4z - 2(x - y - z) - (2y - (x - y - z) - 2z) = 8
 \end{array}$$

Hiervoor zijn zoals bekend speciale procedures ontwikkeld, bijvoorbeeld eerst ‘haakjes wegwerken’ en vervolgens de zo verkregen drie vergelijkingen met drie onbekenden op een bepaalde manier reduceren tot twee vergelijkingen met twee onbekenden, enz. Maar het kan ook ‘slimmer’ worden aangepakt, door iemand die door *inspectie* van de eerste twee vergelijkingen meteen inziet dat $(x - y - z)$ gelijk is aan 2 en $(2y - (x - y - z) - 2z)$ gelijk is aan 4 vervolgens deze waarden *substitueert* in de derde vergelijking. Het laatste kan gewoon ‘uit het hoofd’ gedaan worden, met het snelle resultaat dat $4z - 4 - 4 = 8$, dus dat $z = 4$. We zouden dit kunnen zien als een voorbeeld van ‘natuurlijke intelligentie’. Substitueren in de algebra schijnt namelijk op zichzelf al lastig te zijn. Als we Freudenthal moeten geloven zijn er mensen die dit echt niet kunnen leren – ‘geheel niet toegankelijk zijn voor de methode van formele substitutie, zonder welke geen wiskunde denkbaar is’ – maar waar hij dat vandaan haalde weet ik niet.⁸

Even terug naar computers: oorspronkelijk werd ook gedacht dat zij ingeklede vergelijkingen moesten kunnen omzetten in algebraïsche vergelijkingen,⁹ maar daar is men alras van afgestapt. Computers zover

⁸ Vgl. *Kan het wiskundeonderwijs tot de opvoeding van het denkvermogen bijdragen? Discussie tussen T. Ehrenfest-Afanassjewa en Prof. Dr H. Freudenthal*. Purmerend: J. Muusses, 1951, p. 14. Overigens was Freudenthal wel van mening dat het hanteren van de algebraïsche taal een automatisme moet worden, willen leerlingen verder komen in de wiskunde. Deze les hebben de voorstanders van de zgn. “realistische wiskunde” niet voldoende ter harte genomen. Vgl. Hans Freudenthal, *Logical analysis and critical survey*. In: Hans Freudenthal (ed.), *Report on the Relations between Arithmetic and Algebra*. Groningen: J. B. Wolters, 1962, pp. 20-41.

⁹ Zie H. Brandt Corstius, *Exercises in Computational Linguistics*. Amsterdam: Mathematisch Centrum, 1970. Een fraai voorbeeld dat door deze auteur is uitgewerkt is het volgende probleem: Zoek drie opeenvolgende getallen, waarvan het produkt 21 maal zo groot is als de som. Deze opgave kan

programmeren dat zij een natuurlijke taal beheersen is ondoenbaar gebleken en wordt ook niet meer als een maatgevend probleem voor de filosofie van de AI gezien. U kent de moeilijkheden met zgn. spraakcomputers, zoals die van de inlichtingendienst van de KPN. Maar gelukkig zijn er vooralsnog nog voldoende wiskundige problemen, en ook spelletjes zoals schaken, dammen en niet in de laatste plaats go, waarmee filosofen van de AI hun voordeel kunnen doen in de zin dat zij mogelijke probleemoplossingen door mensen en idem door computers kunnen aanvoeren.

Inmiddels hebt u met uw natuurlijke intelligentie hopelijk al begrepen dat het probleem van Euler ook door terugwerken kan worden opgelost:¹⁰

$$\begin{array}{r} 8 \quad 8 \quad 8 \\ 4 \quad 4 \quad 16 \\ 2 \quad 14 \quad 8 \\ 13 \quad 7 \quad 4 \end{array}$$

Ik zei al dat er nu uitsluitend naar het resultaat kan worden gekeken, in dit geval de oplossing 13, 7, 4, maar ook naar het proces waarin deze oplossing bereikt wordt. Welnu, filosofen van de AI die alleen op het resultaat letten, ik noem ze neo-mechanicisten, zullen bij wijze van spreken al van intelligent gedrag van een computer kunnen spreken als deze die oplossing heeft gegeven, want in dat geval is er door de computer een probleem opgelost waar mensen een zekere intelligentie voor nodig hebben. Neo-intuitionisten, zoals ik hun tegenstanders noem, zullen daar daarentegen niet van overtuigd zijn: zij hechten een grote, zo niet exclusieve waarde aan de gevolgde procedure. Als een computer bijvoorbeeld een oplossing met of voornamelijk met 'brute kracht' vindt, spreken zij niet van 'intelligent' *gedrag*. Dit slaat volgens hen onder meer op de beruchte schaakcomputers die van Kasparov hebben gewonnen.

Het gegeven voorbeeld is overigens ook geschikt om de alvast herformuleerde vraag L3* te illustreren. Mensen kunnen de keuze maken om de weg van het opstellen en oplossen van vergelijkingen in te slaan, bijvoorbeeld omdat zij willen zien of zij dat nog kunnen. Dit is niet geheel

zowel zuiver rekenkundig als algebraïsch worden opgelost, in het eerste geval door regelmaat te ontdekken in de uitkomsten van resp. $1 \times 2 \times 3 : (1 + 2 + 3)$, $2 \times 3 \times 4 : (2 + 3 + 4)$, $3 \times 4 \times 5 : (3 + 4 + 5)$, $4 \times 5 \times 6 : (4 + 5 + 6)$ en $5 \times 6 \times 7 : (5 + 6 + 7)$, geschreven als $3 : 3$, $8 : 3$, $15 : 3$, $24 : 3$ en $35 : 3$ en daarmee de oplossing te voorspellen. Dat de algebraïsche uitwerking voor beginners bijzonder interessant is, kunt u zelf nagaan. Dit soort opgaven kan zeer motiverend zijn voor leerlingen met een sluimerende wiskunde-aanleg!

¹⁰ Vgl. Leonhard Euler, *Vollständige Anleitung zur Algebra*. Neue Ausgabe. Leipzig: Philipp Reclam jun., z.j., p 243: So schwer diese Aufgabe auch zu sein scheint, so zeigt es sich doch, daß sie sogar ohne Algebra aufgelöst werden kann. Man gehe nur in Betrachtung derselben rückwärts.

onbelangrijk als we naar het huidige wiskundeonderwijs in Nederland kijken, als treffend voorbeeld van wat mensen niet meer kunnen doen, wat voor een deel te wijten lijkt aan de filosofie – van de koude grond – van de voorstanders van ‘realistische wiskunde met IT’.¹¹

Enige uitleg bij de benamingen ‘neo-mechanicisten’ en ‘neo-intuitionisten’ is wel op zijn plaats. Daarvoor doe ik even een beroep op uw historische filosofische kennis. U weet dat Kant de stelling verdedigde dat – eenvoudig geformuleerd – menselijk ruimtelijk inzicht en tijdsbesef ten grondslag liggen aan wiskundige kennis. Dit werd later door intuïtionistische wiskundigen zoals L. E. J. Brouwer beperkt tot ‘de tijdsintuïtie’,¹² terwijl Peirce in navolging van de neo-Kantiaan Lange het ruimtebesef fundamenteel achtte.¹³ Peirce sprak zich ook uit over computers *avant la lettre*, de ‘logische machines’ van Jevons en Marquand,¹⁴ vandaar dat ik mij gerechtigd voel hun hedendaagse geestverwanten die met digitale computers te maken hebben ‘neo-intuitionisten’ te noemen. In de tijd van Brouwer heetten zijn tegenstanders in de filosofie van de wiskunde weliswaar ‘formalisten’, maar ik spreek in dit verband liever van ‘mechanicisten’. Deze filosofen van de wiskunde waren met Hilbert van mening dat alle wiskundige problemen beslisbaar zijn.¹⁵ Ook wilden zij geen beroep doen op enig Kantiaans apriori. Volgens sommigen van hen – overigens niet Hilbert – was voor het oplossen van logische en wiskundige problemen geen andere vaardigheid vereist dan het ‘mechanisch’ werken met formules, zodat de genoemde logische machines het werk in principe ook zouden kunnen doen.

Een ander argument tegen zo’n mechanistische visie gaat terug op een beschouwing van William James.¹⁶ Deze zette zich indertijd in de psychologie af tegen een automaten-benadering, omdat hij van mening was dat die geen verantwoording kon geven voor ‘intelligent’ gedrag dat door Brouwer later als ‘teruggaan van doel op middel’ werd beschreven.¹⁷

We komen nu bij de kern van de problematiek van de klassieke filosofie van de AI, of beter de twee kernen ervan: de eerste is de kwestie of het oplossen van wiskundige problemen een bepaald soort inzicht, filosofisch

¹¹ Zie noot 4.

¹² L. E. J. Brouwer, *Over de grondslagen der wiskunde*. Amsterdam en Leipzig: Maas & Van Suchtelen, 1907, p. 81: ... de tijdsintuïtie zonder meer, waarin herhaling als ‘ding in den tijd en nog eens ding’ mogelijk is ...

¹³ Charles Sanders Peirce, *Elements of Logic*. Cambridge, Mass.: Harvard University Press, 1932, p. 40.

¹⁴ C. S. Peirce, Logical Machines. *The American Journal of Psychology*, Vol. I, 1887-1888, pp. 165-170.

¹⁵ David Hilbert, Mathematische Probleme. *Göttinger Nachrichten*, 1900, pp. 253-297. In: David Hilbert, *Gesammelte Abhandlungen. Dritter Band*. Berlin: Julius Springer, 1935, 290-329.

¹⁶ William James, *The Principles of Psychology. Vol. I*. New York: Henry Holt and Company, 1890, p. 8: *The pursuance of future ends and the choice of means for their attainment are thus the mark and criterion of mentality* in a phenomenon. (De cursivering is van James zelf.) De kwestie van de keuze van middelen komt later aan de orde. Vergelijk vraag L3*.

¹⁷ L. E. J. Brouwer, *Over de grondslagen der wiskunde*. Amsterdam en Leipzig: Maas & Van Suchtelen, 1907, p. 82.

‘intuïtie’ genoemd, vereist, de tweede is, of er ook een bepaald soort ‘begrip’ voor nodig is. Het eerste wordt beweerd door neo-intuitionisten als A. D. de Groot en Dreyfus, het laatste door anti-neomechanicisten als Bunge – lang geleden – en Searle, ook al weer een tijdje geleden. Hun argumenten komen aanstonds aan de orde.

Maar laat ik eerst een ietwat kinderachtig voorbeeld geven, om de gedachten te bepalen. Hopelijk kent u het 15-spel niet, want anders is het helemaal flauw.

Er zijn twee spelers, die elk om de beurt een getal uit de verzameling

$$\{1, 2, 3, 4, 5, 6, 7, 8, 9\}$$

moeten kiezen. Degene die het eerst drie, let wel precies drie getallen heeft die samen de som 15 opleveren, heeft gewonnen.

Het blijkt dat mensen nogal wat moeite kunnen hebben om dit spel ‘intelligent’ te spelen, zeker als zij opgevoed zijn met ‘realistisch rekenen met ICT’. We moeten daarom het tijdgebonden karakter van de probleemgerichte benadering niet uit het oog verliezen – vroeger, heel vroeger, zeg in de twaalfde eeuw, werd het als een teken van grote intelligentie gezien dat iemand feilloos getallen onder de honderd kon optellen, aftrekken, vermenigvuldigen en delen, en zo te zien kan dit in de eenentwintigste eeuw in Nederland terugkeren.

Twee spelers uit uw omgeving, we noemen ze A en B, spelen vervolgens het spel, terwijl u het spelverloop hieronder weergeeft:

A:

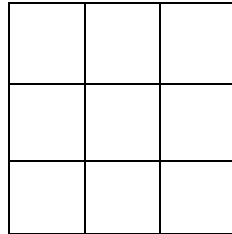
B:

Vanwaar dit 15-spel? Wel, wat zojuist gespeeld is, kan ook anders worden gedemonstreerd, namelijk aan de hand van het bekende, zo niet kinderachtige boter-kaas-en-eieren. De keuzes van de eerste speler komen dan overeen met het plaatsen van een kruisje in een bepaald vakje en de keuzes van de tweede speler met het plaatsen van een rondje, alles volgens een code die ontleend wordt aan een bekend ‘magisch vierkant’ van 3 bij 3:

4	9	2
3	5	7
8	1	6

Dit bevat namelijk alle combinaties van drie verschillende getallen onder de tien met som 15.

Aan de hand van de volgende figuur kunt u nu laten zien welk spel A en B zo-even gespeeld hebben:



Zouden zij deze zetten ook gedaan hebben in deze configuratie? Het ligt niet voor de hand! Het spelen van echte het boter-kaas-eieren is kennelijk eenvoudiger dan het spelen van het 15-spel, vooral ook omdat mensen in het eerste geval de rijen, kolommen en diagonalen zo goed kunnen overzien. Hoewel deze spelen wiskundig niet verschillen, verschillen zij wel ‘psychologisch’. Mensen hebben meer moeite met de rekenkundige versie dan met de ruimtelijke. Voor computers is dit echter andersom. Zij gaan zelfs altijd rekenkundig en nooit ruimtelijk te werk.

Programma’s voor het boter-kaas-en-eierenspel bestaan al zo’n beetje sinds het begin van het computertijdperk.¹⁸ Die programma’s waren natuurlijk ‘goed’ in de zin dat zij fouten van menselijk tegenstanders genadeloos afstraffen en zelf niet konden verliezen. Neo-mechanicisten zullen kunnen zeggen dat zij daarmee al intelligent gedrag vertonen, terwijl die kwestie van het ruimtelijk inzicht dan kennelijk niet essentieel is.

In dit geval was het gemakkelijk om een rekenkundig equivalent te vinden voor een spel waar mensen hun visuele vaardigheden bij kunnen benutten. Neo-intuitionisten zullen daarom een probleem moeten verzinnen waarbij dat niet zo gemakkelijk gaat, of nog beter, waarbij dat zelfs onmogelijk is. In het laatste geval zou er sprake zijn van grenzen aan machinale procedures die er niet zijn voor menselijke intellectuele activiteiten. Daarmee zou vraag L2* beantwoord zijn!

Een eerste kandidaat voor zo’n probleem is het zgn. competitieprobleem. Het gaat hier om het opstellen van een competitieschema voor een even aantal spelers van een spel voor twee personen, zoals schaken en dammen. Naar goed wiskundig gebruik kunnen we dit probleem eerst proberen op te lossen voor vier spelers. Gemakshalve duiden we ze aan met de cijfers 1 t/m 4. Al snel krijgen we de volgende oplossing:

¹⁸ In de NOS-documentaire over de eerste computers in Nederland, *Een heiligdom van buizen en van draad*, uitgezonden op 9 februari 1995, wordt hier een voorbeeld van gegeven.

12 34
13 24
14 23

Vervolgens gaan we over op zes spelers. Dit lukt ook nog wel:

12 34 56
13 25 46
14 26 35
15 24 36
16 23 45

Moeilijker wordt het althans voor ons mensen met tien en meer spelers. U begrijpt het al: een computer zal daar geen moeilijkheden bij ondervinden.

Niettemin hebben computers ook hun grenzen. Bij een experiment, dat uitgevoerd werd door een oud-collega van mij,¹⁹ kwam het programma er aanvankelijk al niet uit bij twintig spelers, vandaar dat de maker ervan enige ‘slimme’ AI-procedures invoerde waarvan de neo-mechanicisten zouden zeggen dat ze de computer ‘kunstmatig intelligent’ maakten. Alleen kwam het gewijzigde programma er bij tweehonderd spelers ook niet meer uit ...

Natuurlijk kunnen we, zeker tegenwoordig, krachtiger computers met meer geheugen inzetten, maar ik geef u de garantie dat er voor elk van hen een aantal is waarbij zij niet meer in staat zijn om een oplossing te genereren. Maar nu komt het: dit geldt niet voor mensen, niet zozeer omdat zij naar believen meer pen en papier kunnen nemen, zoals Peirce ooit opmerkte,²⁰ maar wel omdat zij over kunnen gaan op een andere ‘representatie’. U begrijpt het al, een ruimtelijke representatie.

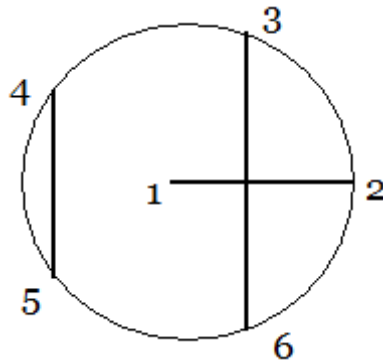
Beeld de spelers op één na af als punten op een cirkel, ongeveer op gelijke afstand van elkaar, en die ene speler als het middelpunt van die cirkel. Neem voor het gemak zes spelers en nummer ze van 1 tot 6. Verbind nu in de figuur 1 met 2, 3 met 6 en 4 met 5. Dit geeft al om te beginnen één ronde van de gevraagde competitie:

12 36 45

Deze is gemakkelijk af te lezen uit de figuur:

¹⁹ Dr Dennis Breuker.

²⁰ C.S. Peirce, Logical Machines. *The American Journal of Psychology*, Vol. I, 1887-1888, p. 169.



Verbind vervolgens 1 met 3 enz.:

13 42 56

.En ga zo door:

14 53 62

15 64 23

16 25 34

Wat voor zes spelers opgaat, gaat ook op voor acht spelers, ja voor elk even aantal spelers, zonder beperkingen.²¹

Het is ook intuïtief duidelijk (!) dat de zo verkregen competitieschema's een regelmatige, 'cyclische' structuur hebben. Neo-intuitionisten zullen dus smullen van dit voorbeeld. Niet alleen omdat er intuïtieve inzichten aan te pas komen, maar ook omdat mensen een procedure hebben gebruikt die vooralsnog aan hen alleen en niet aan computers is voorbehouden, namelijk een overgang, door mij 'transpositie' genoemd, van een cijfermatig, 'numeriek' domein naar een meetkundig, dus ruimtelijk domein. Deze procedure heeft trouwens ook haar vruchten afgeworpen in een tak van de wiskunde die als 'eindige meetkunde' te boek staat.²²

Wat die intuïtieve inzichten betreft: zou het mogelijk zijn dat een computer die het probleem numeriek heeft opgelost voor verschillende aantallen spelers, die oplossingen zo zou kunnen herschrijven dat de cyclische structuur ervan aan het licht komt? En zo ja, zou het programma zelf vervolgens de inductieve generalisatie naar een algemene oplossing kunnen maken? Volgens de neo-intuitionisten ligt de bewijslast hiervoor bij de neo-mechanicisten.

Er is kennelijk iets aan de hand met het zojuist gegeven plaatje van een cirkel met 6 punten dat die intuïtieve inzichten mogelijk maakt. De

²¹ Vgl. Maurice Kraitchik, *Mathematical Recreations*. London: George Allen & Unwin, 1943, p. 226.

²² Zie mijn *Transpositions*.

bekende filosoof van de AI Hubert Dreyfus²³ zou hier in navolging van Wittgenstein²⁴ terecht spreken van een ‘overzichtelijke representatie’.²⁵

Dreyfus gebruikte hierbij overigens niet zoals ik de misschien beladen term ‘intuïtie’, ondanks zijn bekendheid met het werk van A. D. de Groot. Deze scherpzinnige psycholoog had al in zijn dissertatie²⁶ een analyse gegeven van intuïtieve ‘sprongen van doel op middel’ die niet alleen voor de filosofie van de AI van grote waarde is gebleken.²⁷

Later gaf A.D. de Groot met zoveel woorden de volgende stellingen:

- (1) zodra we weten met welke kennis en procedures een probleem is opgelost, kunnen we die kennis en procedures programmeren²⁸
- (2) bij intuïtieve gevolgtrekkingen waarmee een probleem wordt opgelost zijn de eventuele kennis en procedures die daar aan ten grondslag zouden kunnen liggen zelfs niet toegankelijk voor de probleemoplosser, *ergo* intuïtieve gevolgtrekkingen kunnen niet worden geprogrammeerd.²⁹

Een neo-mechanicist hoeft hier niet meteen ondersteboven van te raken, niet dat hij toch intuïtieve gevolgtrekkingen zal willen nabouwen, dat kan niet, maar hij zal zo mogelijk (andere) kennis en procedures introduceren waarmee toch ook hetzelfde resultaat wordt bereikt. Hij ontmoet daarbij echter nog een andere *bottleneck* dan de nu besproken ‘ruimtelijke intuïtie’; er is namelijk ook nog zoets als intuïtief inzicht in oneindige voortzettingen. Mijn standaardvoorbeeld is het bekende pakje Droste-cacao met een plaatje van een verpleegster die een pakje cacao op een

²³ Hubert L. Dreyfus, *What computers can't do. A critique of artificial reason*. New York etc.: Harper & Row, 1972, p.40.

²⁴ Volgens Wittgenstein bevorderen overzichtelijke representaties ons ‘inzicht’ in de zin dat wij ‘verbanden zien’. Zie Ludwig Wittgenstein, *Philosophische Untersuchungen (Philosophical Investigations)*. Oxford: Basil Blackwell, 1958, p. 49 (sectie 122): Die übersichtliche Darstellung vermittelt das Verständnis, welches eben darin besteht, daß wir die ‘Zusammenhänge sehen’.

²⁴ Dr Dennis Breuker.

²⁴ C. S. Peirce, Logical Machines. *The American Journal of Psychology*, Vol. I, 1887-1888, p. 169.

²⁴ Vgl. Maurice Kraitchik, *Mathematical Recreations*. London: George Allen & Unwin, 1943, p. 226,.

²⁵ Vgl. o.a. mijn Perspicuous representations in algebra.

²⁶ A. D. de Groot, *Het denken van den schaker. Een experimenteel-psychologische studie*. Amsterdam: Noord-Hollandsche Uitgevers Maatschappij, 1946, p. 243.

²⁷ Vgl. Henk Visser, Significs, Semiotics and Synthesis. A. D. de Groot's Contributions to Significs. *Theory and Psychology* 9, 1999, pp. 67-87.

²⁸ Vgl. A. D. de Groot, De programmering van het creatieve. In: *Mens en computer. Automatie, industriële en culturele revolutie*. Utrecht en Antwerpen: Aula Boeken, 1963, p.162.

²⁹ Een van A. D. de Groots meest recente formuleringen van deze stelling luidt als volgt: If a programmer has no idea of how a claimed human capability functions, he cannot program it. In the case of intuition, even the person using it does not know how it works. Suppose that psychologists cannot explain it either, we can say: The answer to the question “what computers can't do” is simple – “intuitive processing.” Zie Adriaan D. de Groor & Fernand Gobet, *Perception and Memory in Chess. Studies in the Heuristics of the Professional Eye*. Assen: Van Gorcum, 1996, p. 269.

dienblaadje heeft, waar een verpleegster op is afgebeeld die net zo'n pakje cacao op een dienblaadje heeft.

Ik zal hier niet verder op in gaan. Waar het in de filosofie van de AI om gaat, is dat problemen worden aangegeven die dankzij dit inzicht, en met daaraan ontleende procedures, worden opgelost, waarna neo-mechanicisten maar moeten aantonen dat die problemen ook daarzonder kunnen worden opgelost. Bekend zijn de paradoxen van Zeno, maar voor het volgende probleempje zou ik toch ook al wel een computeroplossing willen zien:³⁰

Iemand heeft 1 gulden. Hij gooit deze in het water en krijgt er 2 van de bank. Daarvan gooit hij er weer 1 in het water en krijgt er daarna 2 van de bank. De guldens gooit hij in het water in dezelfde volgorde, waarin hij ze ontvangt. Als hij dit proces oneindig vaak voortzet, kan hij dan oneindig veel guldens verkrijgen? (Met oneindig vaak is aftelbaar oneindig vaak bedoeld.)

Terug naar A.D. de Groot. Deze heeft zich ook uitgesproken over antwoorden op vraag L3*. Hij deed dit in een klassieke studie, 'De programmering van het creatieve'. Na een aantal voorbeelden van programmeerbaarheid te hebben gegeven, merkte hij namelijk het volgende op – als eerbetoon aan de auteur geef ik een wat uitgebreider citaat:³¹

Minder pertinent moet ik zijn bij de volgende principiële vraag: kan men een creatief proces opvatten als een reeks keuzen uit een telkens beperkt en vooraf bekend aantal handelingsalternatieven? Het ziet er naar uit dat dit voor programmeerbaarheid toch wel noodzakelijk is – zij het dat er altijd de mogelijkheid is dat de alternatieven niet klaar liggen maar ad hoc gegenereerd worden. Hoe dan ook: men moet een boom van vertakkingen hebben of kunnen maken.

Newell en Simon menen dat dit zo is, dus dat er bij het scheppen óók steeds sprake is van kiezen. Psychologisch ben ik het niet met hen eens. Een schilder die 'een keuze doet' uit materialen, uit de mogelijkheden van verf, kleur, manier, vlakverdeling, van richting, dikte en breedte van een penseelstreek enz., is terwijl hij zijn 'keuze' doet, op dat ogenblik psychologisch niet aan het 'kiezen'. Dit laatste vooronderstelt dat een bepaalde verzameling van mogelijkheden voor hem gegeven is, dat hij

³⁰ P. G. J. Vredenduin, *Vijfentachtig wiskundige puzzels met oplossingen*. Groningen: P. Noordhoff, 1964, p. 20 en 52. Zie ook P. G. J. Vredenduin, *Recreatie. Euclides 36, 1960-1961*, p.207 en p. 255.

³¹ A. D. de Groot, *De programmering van het creatieve*. In: *Mens en computer. Automatie, industriële en culturele revolutie*. Utrecht en Antwerpen: Aula Boeken, 1963, p.171.

weet en zich bewust is welke alternatieven er zijn – en dat is psychologisch stellig niet altijd het geval. Aan te nemen dat dit onderbewust wel zo zou zijn, is psychologisch gezien een uitvlucht als men daar verder niets van kan aantonen. Ik blijf er in ieder geval bij dat het verschil maakt of men iets in de (creatieve) situatie ‘vindt’ of dat men ‘kiest’ uit een gegeven stel gevallen.

Terzijde merk ik op dat het hier gemaakte onderscheid tussen ‘een keuze doen’ en ‘kiezen’ doet denken aan Wittgensteins onderscheid tussen ‘een regel volgen’ en ‘volgens een regel te werk gaan’.³²

Een ander onderwerp dat bij keuzes aan de orde is, is het vraagstuk van ‘de vrije wil’ dat zelfs expliciet door Feigenbaum and Feldman in L3 was opgenomen. Deze kwestie is actueel geworden sinds Braitenberg in zijn boek *Vehicles* beweerde een robot, Vehicle 12 genoemd, te kunnen ontwerpen die onvoorspelbaar gedrag vertoont, ondanks het feit dat het mechanisme ervan bekend is. Volgens Braitenberg is dit het bewijs dat zo’n robot een vrije wil heeft. Ook hier volsta ik met een citaat:

I know what the philosophers will reply. They will say that although this may look like free will, in fact it is not. What they have in mind when they use that term is the real power of decision, a force outside any mechanical explanation, an agent that is actually destroyed by the very attempt to put it into a physical frame.

Op zijn beurt beantwoordde Braitenberg dit met de opmerking dat mensen, maar ook dieren en robotjes, althans zijn ‘vehikels’, zelf niet in staat zijn te voorspellen wat er op het volgende moment in hun brein zal gebeuren; hierdoor krijgt zo’n individu het gevoel dat zijn handelingen niet causaal gedetermineerd zijn.³³ U begrijpt dat filosofen dit niet op zich kunnen laten zitten. Het betekent wel dat het kunstmatige intelligentieonderzoek hen weer aan het werk heeft gezet, zelfs met zo’n antiek probleem als dat van de vrije wil.

Tot slot wil ik nog, zij het kort, het tweede kernvraagstuk van de filosofie van de AI bespreken. Anti-mechanicisten huldigen zoals gezegd de opvatting dat er voor het oplossen van wiskundige problemen ‘begrip’ nodig is waar computers niet over kunnen beschikken. Om dit toe te lichten gaan we terug naar de formalisten uit het begin van de twintigste eeuw. Zij zagen wiskunde als een activiteit met symbolen –

³² Zie om te beginnen Ludwig Wittgenstein, *Philosophische Untersuchungen (Philosophical Investigations)*. Oxford: Basil Blackwell, 1958, pp. 38-39 (sectie 82).

³³ Valentino Braitenberg, *Vehicles. Experiments in Synthetic Psychology*. Cambridge, Mass. and London: The MIT Press, 1984, p. 69.

‘symboolmanipulatie’ zouden psychologen later zeggen.³⁴ Daar doet een computer inderdaad aan, alleen zijn zijn symbolen materiële objecten, die wel ‘uitdrukkingen van gedachten’, maar zelf geen gedachten zijn, aldus de antimechanisten, of zo u wilt, neo-mentalisten, om nog maar een ideaaltype te noemen. Hun conclusie ligt voor de hand: computers voeren geen wiskundige berekeningen uit, maar alleen fysische operaties die *wij* – mensen – coördineren met wiskundige bewerkingen. Ik citeer uit een klassieke verhandeling van Mario Bunge:³⁵

Machines, let us repeat, do not add pure numbers; they add turns of cogwheels, electric pulses, etc., which combine in accordance with physical laws peculiar to each such material objects. Machines are not, as our nervous system is, multilevel structures; they cannot retranslate concrete objects into their abstract correlates or vice versa. It is we who perform such a translation, when building and using the machine.

Dit lijkt triviaal, maar filosofisch is er meer aan de hand: hoe komt het dat mensen in staat zijn de stap van het ‘concrete’ naar het ‘abstracte’ te zetten? Volgens Bunge hebben mensen de mogelijkheid een ‘mentaal’ niveau boven het fysiologische op te bouwen, een mogelijkheid die hun geboden wordt door de aanwezigheid van de cortex. Later drukte de filosoof John Searle zich iets simpeler uit toen schreef dat een computer geen idee heeft dat ‘4’ 4 betekent of dat het überhaupt iets betekent, terwijl hij vervolgens ‘intentionaliteit’ een biologisch verschijnsel noemde.³⁶

Daarmee is *een* antwoord gegeven op vraag L4*. Niettemin zou ik nog wel een wetenschappelijke onderbouwing van de laatste stelling willen hebben, terwijl filosofen zich toch ook nog wat ongemakkelijk kunnen voelen bij het horen van de termen ‘het concrete’ en ‘het abstracte’.

Maar er is ook nog een meer laag-bij-de-grond-aspect aan Bunge’s stelling dat *wij* het zijn die de interpretatie voltrekken. Ik geef een voorbeeld. Stel, we hebben een computerprogramma dat concludeert dat *De geschiedenis van mijn kaalheid* van een zekere Marek van der Jagt in

³⁴ Het woord “symbolen” suggereert nog dat er iets gesymboliseerd wordt, maar wanneer de eerder genoemde Hilbert op latere leeftijd van “tekens” sprak, zou elke associatie met iets waarvoor een teken staat, achterwege moeten blijven. Beroemd, of beter berucht is zijn uitspraak *am Anfang ist das Zeichen*, gevolgd door zogenaamde verklaringen van getallen: *Das Zeichen 1 ist eine Zahl. Ein Zeichen, das mit 1 endigt, so daß dazwischen auf 1 immer + und auf + immer 1 folgt, ist ebenfalls eine Zahl, z.B. die Zeichen 1 + 1, 1 + 1 + 1. Diese Zahlzeichen, die Zahlen sind und die Zahlen vollständig ausmachen, sind selbst Gegenstand unserer Betrachtung, haben sonst keinerlei Bedeutung.* (Deze uitspraken deed Hilbert in een voordracht van 1922. Zie wederom zijn *Gesammelte Abhandlungen. Dritter Band*. Berlin: Julius Springer, 1935, p. 163.) Zo blijkt maar weer dat grote wiskundigen geen grote filosofen hoeven te zijn.

³⁵ Mario Bunge, Do computers think? *The British Journal for the Philosophy of Science* 7, 1956, p. 145.

³⁶ Vgl. John R. Searle, Minds, Brains and Programs. *The Behavioral and Brain Sciences*, 3, 1980.

feite door Aron Grunberg is geschreven, of een neurale netwerk dat tot de conclusie komt dat de Japanse zonnebloemen inderdaad door Vincent van Gogh geschilderd zijn. Moeten wij deze conclusies dan ook maar aanvaarden? Noch het computerprogramma, noch het neurale netwerk heeft enig begrip van het object waar hij of het zich mee bezig heeft gehouden. De computer ‘weet’ niet dat het om een psychologische roman gaat, het neurale netwerk niet dat het om een schilderij gaat – vergelijk vraag L1* – en toch komen zij tot conclusies die door ons zo geïnterpreteerd worden dat zij van toepassing zijn op artistieke producten.

Waarom zouden we die conclusies *filosofisch gesproken* – praktisch is het een andere zaak – overnemen? Zou iemand niet een boek in de stijl van Grunberg kunnen hebben geschreven en iemand anders een schilderij in de stijl van Van Gogh kunnen hebben gemaakt? Sterker nog: er zijn mensen die weten in de sterkste zin van het woord dat *De geschiedenis van mijn kaalheid* door Grunberg is geschreven of wisten dat de Japanse zonnebloemen door Van Gogh zijn geschilderd, namelijk Aron Grunberg en Vincent van Gogh zelf. Zij hebben dat hele computerprogramma respectievelijk dat hele neurale netwerk dus niet nodig. Is niet ook hun ‘begrip’, meer nog dan kunstenaars-in-de-stijl-van niet *van een andere orde* dan de ‘kennis’ die een computerprogramma of een neurale netwerk heeft van de structuur van een compositie? Denk weer aan vraag L1*. En is die structuur en het besef ervan niet ten dele intuïtief tot stand gekomen? Denk aan vraag L2*. Leste Grunberg trouwens het probleem wie *De geschiedenis van mijn kaalheid* schreef niet op (nou ja) met dat befaamde bewustzijn waar sommige filosofen het zo druk mee lijken te hebben?

Vooralsnog lijken de neo-intuitionisten het pleit te hebben gewonnen – alleen maken zij in hun argumentatie geen gebruik van wetenschappelijke resultaten en zijn er nog heel wat onduidelijkheden in de gebezigde noties. Het is mijn overtuiging dat een grondige wetenschappelijke analyse en beschrijving van probleemoplossingen uitkomst kan bieden. Het is gebleken dat daarbij, althans voor wiskundige problemen, kennis en procedures worden ontdekt die vervolgens (indachtig A. D. de Groot’s eerste stelling) kunnen worden geprogrammeerd. Ook over ‘intuïtieve inzichten’ is het laatste woord nog niet gezegd. Het lijkt zinvol daar tenminste twee aspecten aan te onderscheiden, (1) ‘intuïtieve gevolgtrekkingen’ (*intuitive inferences*) en (2) ‘veelbelovende conclusies’ (*promising conclusions*).³⁷ Hierdoor worden problemen met intuïtie beter bespreekbaar, waar neo-mechanicisten hoop uit kunnen putten: zouden computerprogramma’s door middel van (1*) analogieredeneringen en (2*) inductieve en

³⁷ Zie o.a. mijn Squangles. Intuitive inferences and promising conclusions.

deductieve redeneringen niet ook oplossingen van problemen kunnen vinden die sommige mensen intuïtief inzien? Maar hierover wellicht een volgende keer!