

# MENSELIJKE EN KUNSTMATIGE INTELLIGENTIE

## Numerieke intelligentie

Er is een tijd geweest dat mensen die met papier en potlood goed konden rekenen buitengewoon intelligent werden geacht, ten eerste omdat zij dat kennelijk hadden konden leren, en ten tweede vanwege hun prestaties op dat gebied, omdat die voor weinigen waren weggelegd. In de Middeleeuwen waren er althans niet zoveel jongeren die naar een Universiteit gingen, want dat was de plaats waar je zulk rekenen kon leren, wanneer je daar tenminste aanleg voor had. Hierbij moet worden opgemerkt dat bijvoorbeeld vermenigvuldigen van twee getallen van zeg drie cijfers anders gebeurde dan in latere tijd, zeg de vorige eeuw, om maar te zwijgen van de eenentwintigste eeuw.<sup>1</sup>

Eén van de manieren, waarop in de zestiende eeuw de vermenigvuldiging  $934 \times 314$  werd uitgevoerd, staat hieronder, gevolgd door een rekenwijze die ik zelf nog geleerd heb:

	9	3	4	
2	2	0	i	
9	0	9	0	3
3	3	3	0	i
	3	i	i	4
	2	7	6	

<sup>1</sup>Naar Frank Swetz, *Capitalism and Arithmetic. The New Math of the Fifteenth Century*. Chicago: Open Court, 1986.

$$\begin{array}{r}
 934 \\
 \underline{314} \\
 3736 \\
 9340 \\
 \underline{280200} \\
 293276
 \end{array}$$

Het feit dat het enige tijd kost om te doorgronden wat hier allemaal plaatsgevonden heeft, duidt er op dat er inderdaad een soort intelligentie voor nodig is om zulke rekenvaardigheden te kunnen *aanleren* en in het verlengde hiervan zo te kunnen *blijven* rekenen. Deze vorm van intelligentie noem ik in navolging van *Psychologie Magazine* ‘numerieke intelligentie’.

Weliswaar leert de jeugd van tegenwoordig niet of nauwelijks meer zulke ‘ingewikkelde’ berekeningen te maken, want zij krijgen bij voorkeur sommetjes die ze zonder veel tussenstappen kunnen oplossen. En dat komt goed uit, want langere berekeningen schijnen niet te kunnen worden gecitotoetst. Maar dat wil niet zeggen dat het merendeel van de jeugd van tegenwoordig dat niet meer zou *kunnen* leren.

Er zit echter meer achter: het idee bestaat dat alleen maar bewerkingen mogen worden aangeleerd waarvan de leerlingen *begrijpen* waarom ze tot het gestelde doel voeren. Dat was bij de zestiende-eeuwse methode waarschijnlijk al niet het geval, en voorzover dat ooit het geval was bij de tweede methode dacht geen sterveling er bij de uitvoering meer aan waarom dat allemaal zo gedaan werd. Dit gold des te meer voor de zogenaamde staartdelingen, die nu in Nederland helemaal afgeschaft zijn. In plaats ervan is een maniertje gekomen waarmee alleen delingen met relatief kleine getallen kunnen worden uitgevoerd. Er is hier ten lande kennelijk een haat ontstaan tegen alle ‘automatismen’, d.w.z. het zonder nadenken uitvoeren van bewerkingen waarvan niet of niet meer duidelijk is waarom ze zo en niet anders moesten worden uitgevoerd. Dat werd kennelijk door een nieuwe generatie van zelfbenoemde ‘onderwijskundigen’ niet als een ‘intelligente’ bezigheid gezien. Dat leerlingen vroeger een geluksgevoel kregen als hun deling op 0 uitkwam, is kennelijk ook niet tot hen doorgedrongen:

$$\begin{array}{r}
 314 / 293276 \setminus 934 \\
 \underline{2826} \\
 1067 \\
 \underline{942} \\
 1256 \\
 \underline{1256} \\
 0
 \end{array}$$

Het zal duidelijk zijn dat dit onderwijskundige novum alleen maar kon worden ingevoerd omdat de grotere berekeningen aan de handcomputer konden worden overgelaten. En juist daardoor kon het lijken alsof zulke berekeningen geen intelligent werk meer zouden zijn. Dit lijkt op de situatie van de jaren zestig toen de groeiende hoeveelheid giro-overschrijvingen de pan uit sprongen en de mogelijkheden van automatisering werden onderzocht. Aanvankelijk vergden de overschrijvingen hoe langer hoe meer menskracht, maar werd het werk als zodanig niet in strijd met de ‘menselijke waardigheid’ bevonden.<sup>2</sup> Later, toen computers het werk overgenomen hadden, werd met verbazing teruggekeken naar de tijd dat mensen het deden: dat kon toch niet meer, de hele dag dat tijdrovende en geestdodende werk met zoveel ‘personeel’ – want denk maar niet dat de chefs zich daarmee bezighielden! Maar wil dat zeggen dat er geen intelligentie voor nodig was? Dat zou een belediging voor die harde werkers geweest zijn.

Mijn standpunt is dus duidelijk: er bestaat zoiets als numerieke intelligentie op grond waarvan mensen ingewikkelde berekeningen met potlood en papier kunnen maken. En dat heeft consequenties voor de wijze waarop we over computers kunnen spreken: als zij inderdaad berekeningen kunnen uitvoeren waar mensen numerieke intelligentie voor nodig hebben, dan kunnen we zeggen dat die computers in dat geval ‘kunstmatige intelligentie’ vertonen. Dit is in overeenstemming met de volgende definitie:

The capability of a device to perform functions that are normally associated with human intelligence

(De mogelijkheid van apparatuur om taken te vervullen die normaliter geassocieerd worden met menselijke intelligentie)

Bij zulke taken denken we in eerste instantie aan probleemoplossen, maar vervolgens ook aan ‘concrete’ activiteiten, variërend van meet- en regeltechnische tot gerobotiseerde processen. Een waarschuwing is hier op zijn plaats: het is niet zo dat zulke kunstmatige intelligentie dezelfde vaardigheden vereist of zou moeten vereisen als die waar mensen *de facto* problemen mee oplossen, integendeel zou ik bijna zeggen. Als deze suggestie uitgaat van het gebruik van de term ‘kunstmatig’, dan kan beter worden gesproken van ‘automatische intelligentie’, ‘computationele intelligentie’, ‘computerintelligentie’, ‘digitale intelligentie’ of wat dies meer zij.

Toch wil dat wil niet zeggen dat de procedures waarmee computers problemen oplossen überhaupt geen tegenhangers kunnen hebben van

---

<sup>2</sup> Vergelijk de documentaire *Een heiligdom van buizen en van draad* (NOS 14.02.1995).

vaardigheden waarvoor mensen intelligent genoeg moeten zijn. Om een voorbeeld te geven: als er een wiskundig probleem is waar geen algemene oplossing voor bestaat, kan het wenselijk zijn om in concrete gevallen met ‘systematisch uitproberen’ te onderzoeken of er een oplossing is. Welnu, om dat te kunnen doen is óók intelligentie vereist, denken we maar terug aan opgaven als het uitvoeren van grote vermenigvuldigingen of *dito* staartdelingen. Alleen hebben wiskundigen daar niet zoveel plezier aan en zoeken zij liever naar formules waarmee in één klap een hele klasse van problemen mee kan worden opgelost. De neiging bestaat om een aanpak van een probleem die voornamelijk uit ‘systematisch uitproberen’ bestaat, als een niet bijster of een in het geheel niet intelligente oplossingswijze te beschouwen en in het verlengde daarvan een computerprogramma dat voornamelijk op ‘brute kracht’ is gebaseerd ook maar niet als, zij het kunstmatig, intelligent te beschouwen. Ik doe dat dus niet (meer) en onderschrijf dan ook de volgende stelling:

De wijze waarop computers problemen oplossen is irrelevant voor de beoordeling of zij kunstmatig intelligent zijn

### Relevante vragen

Met deze laatste stelling en de gegeven definitie van kunstmatige intelligentie zijn heel wat schijnproblemen van het type ‘kunnen machines denken?’ of ‘kunnen computers intelligent zijn?’ uit de wereld geholpen en kunnen we ons wijden aan de serieuze vragen over computers, zoals:

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?

Dit zijn mijn vrije vertalingen van de vier *kennistheoretische* (‘epistemologische’) vragen die Marvin Minsky ooit geformuleerd heeft

voor zijn beredeneerde bibliografie van de kunstmatige intelligentie.<sup>3</sup> Maar er is niets tegen om deze vragen ook als ethische problemen te zien: zijn er *morele* grenzen aan machinale verrichtingen? Het zou te ver voeren om hier elk van deze vragen te behandelen, en daarom beperk ik mij tot vragen van de vorm:

Welke mogelijkheden en onmogelijkheden, respectievelijk wenselijkheden en onwenselijkheden zijn er voor computers om bepaalde problemen op te lossen?

Met behulp van zulke vragen kan gemakkelijk een overzicht van de geschiedenis van de kunstmatige intelligentie worden geschreven. Wie herinnert zich niet de vraag of een schaakcomputer van een grootmeester zou kunnen winnen? Deze vraag werd indertijd voorafgegaan door de vraag of een schaakcomputerprogramma zijn programmeur zou kunnen verslaan.<sup>4</sup> Het antwoord op beide vragen is bevestigend gebleken, al heeft het enige tijd geduurd voor de wereldkampioen tegen een schaakcomputer het loodje moest leggen. Meer vragen van soortgelijke aard waren indertijd, d.w.z. voor 1960, ‘Hoeveel kunnen machines leren?’<sup>5</sup>, ‘Kan vertalen gemechaniseerd worden?’<sup>6</sup>, ‘Zou een machine waarschijnlijkheidsoordelen kunnen vellen?’<sup>7</sup>, ‘In hoeverre kan de administratie worden gemechaniseerd?’<sup>8</sup> en ‘Kan inductie gemechaniseerd worden?’<sup>9</sup>

Van meer recente datum zijn vragen als, ‘Kan een computer winnen van een go-speler negende dan?’<sup>10</sup> ‘Kunnen computers perfect Nederlands leren?’<sup>11</sup> ‘Kunnen computers contracten maken?’<sup>12</sup> en ‘Kunnen computers rechtspreken?’<sup>13</sup> Zullen de eerste twee vragen geen negatief antwoord krijgen, terwijl de laatste vragen meteen de vraag oproepen of computers dat wel ‘mogen’? Jos Baeten en Tijn Borghuis van de TUE hebben hun

---

<sup>3</sup> In: Edward A. Feigenbaum and Julian Feldman (eds.), *Computers and Thought*. New York etc.: McGraw-Hill, 1963, pp.471-472.

<sup>4</sup> W. R. Ashby, Can a mechanical chess player outplay its designer? *British Journal for the Philosophy of Science*, 3, 1951, pp. 44-57.

<sup>5</sup> A. D. Booth, How much can machines learn? *Automatic Data Processing*, 2, 1960, pp. 22-26.

<sup>6</sup> Y. Bar-Hillel, Can translation be mechanized? *American Scientist*, 42, 1954, pp. 248-260.

<sup>7</sup> Good, I. J. Could a machine make probability judgments? *Computers and Automation*, 8, 1959, pp. 14-16, pp. 24.-26.

<sup>8</sup> J. Merriman, D. W. G. Wass and S. Gill, To what extent can administration be mechanized?. In: D. V. Blake and A. M. Uttley (eds.), *Proceedings of the Symposium on Mechanization of Thought Processes*. London: H. M. Stationary Office, 1959.

<sup>9</sup> V. Somenzi, Can induction be mechanized? In: C. Cherry (ed.), *Information theory. Proceedings of the 3d London Symposium on Information Theory*. London: Butterworth; New York: Academic Press, 1956.

<sup>10</sup> Thomas Caldwell.

<sup>11</sup> Rudy Beernink.

<sup>12</sup> Robin Widdison.

<sup>13</sup> H. J. van den Herik, *Kunnen computers rechtspreken?* Arnhem: Gouda Quint, 1991.

antwoord hierop al klaar: 'computers *zullen* niet rechtspreken'.<sup>14</sup> En Lambèr Royakkers van de TUE bracht het als volgt onder woorden:

Over de vraag 'kunnen computers rechtspreken?' heeft Royakkers zelf een uitgesproken mening. "De vraagstelling alleen al is een teken van de waanzin onzer tijd. Hoe intelligent de computer ook moge worden, hij zal zich altijd verre moeten houden van authentieke menselijke aangelegenheden, zoals rechtspraak, opvoeding en politiek. Een zinvolle vraag zou wel kunnen zijn: 'In hoeverre kunnen computers ondersteuning verlenen bij juridische beslissingen?'"<sup>15</sup>

Hier valt nogal wat over te zeggen, want wat zijn 'authentiek menselijke aangelegenheden'? Is rekenen dat soms ook niet? En computers kunnen toch onder bepaalde voorwaarden met succes worden ingezet om leerlingen en studenten bepaalde vaardigheden aan te leren? Waar ligt de grenslijn tussen 'ondersteunen', 'begeleiden', 'leiden' en tot iets 'dwingen', of tussen 'adviseren' en 'de wet voorschrijven'? Om een eenvoudig voorbeeld te geven: als er besloten is de constatering van snelheidsoverschrijdingen en de verdere afhandeling ervan geheel aan computers over te laten, *moet* je dan de eventuele boete betalen? Kortom, de vraag of computers bepaalde activiteiten wel *mogen* ondernemen is niet zo een twee drie te beantwoorden.

Een aantal vragen over zaken die computers al dan niet *kunnen* oplossen is in de periode van 1960 tot nu al wel beantwoord. Ik noemde al met zoveel woorden het verlies van Kasparov tegen Deep Blue, maar daar staat tegenover dat computers niet in staat bleken serieuze stellingen van de Euclidische meetkunde te bewijzen. Gelerntner's onderzoeksprogramma liep dood, ondanks de toegenomen 'rekenkracht'.<sup>16</sup> (Hoe zou een computerbewijs van de stelling van de buitenhoek er uit moeten zien?<sup>17</sup>) Wel was van meet aan duidelijk dat met computers successen zijn te behalen in gebieden waar (creatieve) mensen algoritmen voor kunnen ontwikkelen. Zo bereikte al in de jaren zestig van de vorige eeuw Ir. Theo Bruyn onder leiding van E. W. Beth en A. Heyting resultaten op het gebied van de automatische bewijsvoering in de projectieve meetkunde:

---

<sup>14</sup> J. Baeten en T. Borghuis, Computers zullen niet rechtspreken. *Automatisering Gids* van 30 maart 2001.

<sup>15</sup> Gerard Verhoogt.

<sup>16</sup> Vergelijk: Desmond Fearnley-Sander, bespreking van A. Quaife, *Automated Development of Fundamental Mathematical Theories*. Dordrecht: Kluwer Academic Publishers, 1992.

<sup>17</sup> Vergelijk mijn *Aristarchos' visit to Euclid..*

Ir. Th. Bruyn zette zijn onderzoek voort betreffende automatische bewijsvoering in de projectieve meetkunde. (...) In nauwe samenwerking met Ir. Bruyn begon, ook in 1963, Dipl. Math. H. Fangmeyer met het programmeren van de resultaten voor de IBM 7090. Hij is er in geslaagd verschillende van de theoretische algorithmen, door Ir. Bruyn gevonden en omgezet in computer-efficiënte algorithmen, te programmeren.<sup>18</sup>

De IBM 7090 was de eerste commerciële getransistoriseerde computer. Er werden Saturnus-vluchten mee gesimuleerd ten behoeve van het Apolloproject, dat zoals bekend begon met een menselijke tragedie. We krijgen een indruk van de mogelijkheden van zo'n computer als we te weten komen dat een IBM 7090 in 1961 voor het eerst meer dan honderduizend decimalen van  $\pi$  berekende.

Problemen oplossen en problemen verzinnen

Voorzover computers hoe langer hoe meer in staat waren bepaalde complexe problemen op te lossen, was dit inderdaad voor een deel te danken aan hun toegenomen rekenkracht. Er zijn schaakcomputers die zo'n zes miljoen stellingen per seconde 'doorlopen'.<sup>19</sup> Hier is niets 'mis' mee, dat wil zeggen, dit maakt zulke computers niet 'minder' kunstmatig intelligent. Wel heeft de verbreiding van schaakcomputers het correspondentieschaak om begrijpelijke redenen om zeep geholpen ...

Toch komt het voor dat problemen niet met rekenkracht alleen zoals systematisch uitproberen kunnen worden opgelost, bijvoorbeeld:

Zijn er oneindig veel priemgetallen?

U weet het antwoord op deze vraag en kunt dat ook bewijzen. Maar geldt dat ook voor de volgende vraag?

Zijn er oneindig veel priemgetallen van de vorm  $n^2 - n + 1$ ?

Vindt u dit te ver gaan, dan kan ook het voorbeeld van het opstellen van een competitieschema voor een voldoende groot aantal schaakspelers worden gegeven. Op een gegeven moment gaat dit het vermogen van elke computer te boven die alleen maar met 'brute kracht' te werk gaat.

---

<sup>18</sup> Euratom Ontwerp-eindrapport 032-62-121, pp. 1-2. Beth-Archief 189-205. Zie ook: P. van Ulsen, *E. W. Beth als logicus*. Amsterdam: ILLC Dissertation Series 2000-04, 2000, pp. 206-209, met citaten uit E. W. Beth, Opmerkingen over een meetkundige 'redeneermachine', manuscript.

<sup>19</sup> Naar Frans Morsch.

Ik kom hier op terug. Intussen kunnen we opmerken dat het *verzinnen* van nieuwe problemen een vaardigheid is die een stapje verder gaat of beter gezegd voorafgaat aan het *oplossen* ervan. De Nederlands-Amerikaanse computerkundige en ‘computerfilosoof’ Jan Amkreutz achtte dit een bij uitstek menselijk vaardigheid:

(...) as powerful as computers are in solving problems, humans are more powerful yet, in creating new ones. (...) Problems are like puzzles: you have to create them, before you can solve them.<sup>20</sup>

*Significante* problemen onderkennen is iets wat zelfs bij het traditionele rekenonderwijs al kon voorkomen. Als leerlingen van de Lagere School (!) ‘vroeger’ een opgave gemaakt hadden als

$$\frac{6789 \times 6789 - 2345 \times 2345}{9134}$$

en tot hun grote vreugde de staartdeling op 0 hadden zien uitkomen, zouden zij zich hebben kunnen afvragen hoe de ‘meester’ toch van die mooie sommen kon maken. Voorzover zij daar niet achter konden komen, zou hun wiskundeleraar van het ‘voortgezet lagere onderwijs’ of de ‘middelbare school’ het hun hebben verteld. (‘Had ik dat geweten ...’)

Toegegeven, ondanks de capaciteitsbeperkingen van computers zijn er heel wat problemen volledig met behulp van algoritmen opgelost. Ik noem alleen de achtpuzzel, een schuifpuzzel die indertijd een belangrijke plaats innam in een bekend boek van Nils Nilsson,<sup>21</sup> waarvoor inderdaad een complete oplossing voor is gevonden: als een achtpuzzelprobleem een oplossing heeft, bestaat die uit ten hoogste dertig zetten.<sup>22</sup>

Toch is dit niet de enige vooruitgang in de digitale wereld. Ook de programma’s werden verbeterd, bij schaakprogramma’s dankzij de ontwikkeling van speciale zoektechnieken en de implementatie van meer en meer schaakkennis zonder dat dit laatste de zoeksnelheid al te zeer verkleinde. Ik had het zojuist over aantallen zetten voor achtpuzzelproblemen. Natuurlijk komt een voldoende intelligent mens wel uit elk oplosbaar achtpuzzelprobleem, maar hoe kan hij of zij er zeker van zijn dat de oplossing in een minimaal aantal zetten bereikt is? Studenten kunstmatige intelligentie weten dat hier speciale algoritmen voor zijn bedacht, zoals het A\*-algoritme, een teken dat AI-onderzoek ook theoretisch vruchtbaar is geweest. De gevonden methoden bestonden

---

<sup>20</sup> J. Amkreutz, *Digital Spirit. Minding the future*. Bloomington. IN: 1stBooks, 2003, p. 329.

<sup>21</sup> N. Nilsson, *Principles of Artificial Intelligence*. Berlin etc.: Springer-Verlag, 1982, pp. 18- 27 en pp. 72-88.

<sup>22</sup> P. D. A. Schofield, Complete solution of the ‘Eight-puzzle’. In: N. L. Collins and D. Michie (eds.), *Machine Intelligence*. Edinburgh: Oliver and Boyd, 1967, pp. 125-133.



niet in de tijd dat er nog geen computers waren. Zelfs de achtpuzzel is aanleiding geweest tot een verbetering van programma's doordat er methoden bedacht werden om macro-operatoren te genereren.

Programma's die gebruik maken van zulke 'slimme' procedures danken dat vanzelfsprekend aan menselijke programmeurs. Daarom spreek ik in dit geval van *afgeleide* kunstmatige intelligentie.

Dankzij de creativiteit van programmeurs zijn computers in staat optimale oplossingen te genereren. Hun kunstmatige intelligentie is afgeleid van de menselijke intelligentie.

Wat dat betreft heb ik zelf positieve ervaringen. Mijn collega Jeroen Donkers schreef voor mij diverse computerprogramma's waarmee ik nieuwe eindige meetkunden op het spoor kwam.<sup>23</sup> En ongetwijfeld zijn u soortgelijke successen bekend.

### Concrete problemen

Kunstmatige intelligentie werd oorspronkelijk geassocieerd met 'verstandelijke' functies, het oplossen van problemen door rekenen en redeneren. Mensen mochten dat in gedachten of luidop doen, en zonodig gebruik maken van papier en potlood om berekeningen of afbeeldingen te maken, maar verder dan dat ging het niet. Allengs kwamen er echter ook operaties in zicht die een ander dan puur intellectueel karakter hadden, met als vroege spraakmakende voorbeeld computerkunst. Harold Cohen liet als een van de eersten zien dat daar ook kunstmatige intelligentie bij te pas kan komen. Vervolgens werden kunstmatige intelligentie en robotica gecombineerd. Een bekend overzicht van de geschiedenis van de kunstmatige intelligentie noemt het project ALVINN – Autonomous Land Vehicle In a Neural Network – dat in 1989 van start ging.<sup>24</sup>

Eerder werd ook wel een verband gezien tussen kunstmatige intelligentie en 'concrete' problemen, maar de aan de weg timmerende auteurs (McCarthy en Hayes) hielden zich toen uitsluitend bezig met het *formuleren* van 'naieve natuurkunde' – *naive physics* – in de vorm van principes. Terwijl toch allang bekend was dat sommige intelligentieopgaven, anders dan luciferproblemen, alleen met '*fysiek* uitproberen' kunnen worden opgelost. Bijvoorbeeld:

Leg vijf munten zó tegen elkaar dat elke munt wordt aangeraakt door de andere vier.<sup>25</sup>

---

<sup>23</sup> Zie mijn *Finite geometries*.

<sup>24</sup> Naar Bruce G. Buchanan.

<sup>25</sup> *Donald Duck* 23, 2004, p. 43.

Tegenwoordig zijn robots helemáál populair en sommige auteurs beweren dat het daar in feite naar toe gaat bij kunstmatige intelligentie:

Most of Artificial Intelligence will eventually lead to robotics. Most neural networking, natural language processing, image recognition, speech recognition/synthesis research aims at eventually incorporating their technology into the epitome of robotics - the creation of a fully humanoid robot.

(...)

Basically, a robot is a machine designed to do a human job (excluding research robots) that is either tedious, slow or hazardous. It is only relatively recently that robots have started to employ a degree of Artificial Intelligence in their work - many robots required human operators, or precise guidance throughout their missions. Slowly, robots are becoming more and more autonomous.<sup>26</sup>

De opmerking dat robots in eerste aanleg ontworpen worden om mensen vervelend, langzaam of gevaarlijk werk uit handen te nemen, is koren op mijn molen. De positieve resultaten op dit gebied zijn bekend genoeg. Om één enkel voorbeelden te noemen: bij British Telecom worden robots ingezet om milieugevaarlijke stoffen uit verouderde telefoontoestellen te halen. ‘Aan het karwei kunnen geen mensen worden gezet, tenzij ze van top tot teen in beschermende kleding worden gestoken en van zuurstofmaskers gebruik maken. Dat zou de snelheid en nauwkeurigheid van werken nadelig beïnvloeden.’<sup>27</sup>

Maar de feitelijke constatering dat nog veel robots door mensen bediend of geleid moeten worden is ook juist. In geïsoleerde of zelfs zo kunstmatig mogelijke omgevingen is dat misschien minder nodig, maar in de praktijk komen die om begrijpelijke redenen minder voor. Medische operaties vinden nu eenmaal niet plaats op niet-bestaande humanoïden.

Het is daarom van het grootste belang dat de mensen die de robots besturen of in het gareel houden, tenminste terdege op de hoogte zijn van hun capaciteiten, anders kan het gauw misgaan. Dat geldt zeker voor veiligheidskritische apparatuur. Dit probleem doet zich overigens al voor bij niet robotachtige computers en nog zeer recent hebben we hier een voorbeeld van kunnen vernemen: bij de in 2004 gehouden match om het wereldkampioenschap schaken tussen Kramnik en Leko had de eerste met zijn secondanten voor de achtste matchpartij een variant van zevenentwintig zetten voorbereid met een dameoffer op de vijfentwintigste zet die ‘tot groot voordeel’ zou leiden. Als gevolg hiervan

---

<sup>26</sup> Naar James Matthews.

<sup>27</sup> Aldus Nico Baaijens, Robot goed voor mens en milieu. *Algemeen Dagblad*, 20.10.1995.

speelde Kramnik zijn eerste vijftwintig zetten *a tempo*, terwijl Leko al op zijn zestiende zet aan het denken was geslaan. Helaas voor Kramnik deed Leko op zijn vijftwintigste zet niet wat Kramnik en zijn assistenten op grond van de computerresultaten verwacht hadden en het gevolg hiervan was dat hij de partij jammerlijk verloor. Het commentaar van Ger Ligterink in *de Volkskrant* van 9 oktober 2004 na de vijftwintigste zet van Leko liegt er niet om:

Nu begon Kramnik te denken. Te laat, de partij is voorbij want na 26. a7 komt (...) 33. Pg1, Pf2 mat.

Ligterink had er nog even zijn eigen schaakcomputer voor geraadpleegd en daarmee leverde hij een prachtige illustratie van de negatieve versie van mijn stelling over afgeleide kunstmatige intelligentie:

Fritz vindt deze variant natuurlijk ook wel maar het programma heeft tijd nodig. Die was hem door Kramnik en zijn secondanten niet gegeven.

De generalisatie van dit verschijnsel ligt voor de hand:

Kunstmatig intelligente programma's moeten altijd nog worden opgestart en stopgezet door menselijke gebruikers. Maar het is de vraag of hun (menselijke) intelligentie de kunstmatige intelligentie wel optimaal kan benutten.

Tekortkomingen van mens-machinesystemen

Controle op de uitvoering van programma's zoals dat van een schaakcomputer is om begrijpelijke redenen niet meer mogelijk. Vandaar de bekende discussies over de correctheid van computerbewijzen, bijvoorbeeld voor het beroemde vierkleurenprobleem. Toch vertrouwen in het alledaagse leven veel mensen zelfs evident tekortschietende programma's, tenminste als ik het aantal studenten tel die een essay inlevert nadat zij het aan een spellingscontrole hebben onderworpen. Hun valt niets op in de volgende tekst:

Henk Visser verdedigt de stelling dat wij veel te weinig oog hebben voor de negatieve kanten van ons computergebruik. Denk maar eens aan de schrijffouten die door de spellingscontrole heen komen, maar die de schrijver zelf ook niet meer herkent.

Deze twee aspecten van kunstmatige intelligentie verdienen nadere aandacht, ik herhaal:

(1) de controle op de uitvoering van kunstmatig intelligente programma's

(2) het vertrouwen in de uitvoering van kunstmatig intelligente programma's

De *volledige* controle op de uitvoering van kunstmatig intelligente programma's is namelijk in het algemeen niet mogelijk, terwijl het *volledige* vertrouwen in de uitvoering van kunstmatig intelligente programma's niet wenselijk is. Nu is dat niets bijzonders, want hetzelfde geldt voor activiteiten die met menselijke intelligentie worden verricht. Maar de controle en het vertrouwen dáárop blijven meespelen bij de controle en het vertrouwen op kunstmatig intelligente verrichtingen, dus het wordt alleen maar moeilijker. Ik heb het nu niet meteen over zulke flauwe grappen als de ontsnapping van de robot Gaak,<sup>28</sup> maar om serieuze problemen.

Op dit punt gekomen kan ik er niet onder uit even de argumenten te bekijken van Bill Joy, die – vrij vertaald – onderscheid maakte tussen twee scenario's, in het kort:<sup>29</sup>

- (i) mensen behouden de controle over intelligente machines
- (ii) intelligente machines kunnen hun eigen gang gaan

In het eerste geval is het gevaar volgens Joy reëel dat de controle over grote systemen aan een kleine groep experts voorbehouden is, met alle mogelijke gevolgen van dien. Zij zullen zich aan een gedragscode moeten onderwerpen, wil het niet mis gaan:

Verifying compliance will also require that scientists and engineers adopt a strong code of ethical conduct, resembling the Hippocratic oath, and that they have the courage to whistleblow as necessary, even at high personal cost.

Het merkwaardige is dat Joy het eerste scenario nog ter sprake bracht, terwijl het tot hem doordrong dat er grenzen zijn aan de betrouwbaarheid van de software:

My continuing professional work is on improving the reliability of software. Software is a tool, and as a toolbuilder I must struggle with the uses to which the tools I make are put. I have always

---

<sup>28</sup> Naar Dave Higgins, Robot on the run). Vergelijk: Amanda Knuveld, De vlucht. *NRC Handelsblad* 26.06.2002.

<sup>29</sup> B. Joy, Why the future doesn't need us. *Wired* 8.04, April 2004.

believed that making software more reliable, given its many uses, will make the world a safer and better place; if I were to come to believe the opposite, then I would be morally obligated to stop this work. I can now imagine such a day may come.

Hieruit volgt dat het tweede scenario helemaal niet in aanmerking zou komen om in gang te zetten. Maar toch is de complete computerisering van boetes voor te snel rijden al op sommige Nederlandse snelwegen ingevoerd, zij het dat je nog wel bezwaar kan worden aangetekend tegen ‘computerbeschikkingen’ waar geen agent meer aan te pas komt.

Toekomstscenario’s bedenken is aardig, maar er zijn al genoeg problemen in het heden: er is nog zoveel te verbeteren aan veel van de alledaagse bestaande software dat alleen al daardoor de gedachte aan vertrouwen in de uitvoering van kunstmatige intelligente programma’s vergezocht lijkt. Nogmaals, vertrouwen is misplaatst zodra ook maar één van twee minpunten het geval kan zijn, te weten,

(I) het tekortschieten van mensen,

hetzij

- (1) door hun onvermogen om de verrichtingen van de computer na te rekenen, terwijl daar toch fouten in zouden kunnen zitten – zie ook (II) –

of

- (2) doordat zij onvoldoende afgestemd zijn op het werken met computers en zelf fouten maken

en

(II) het tekortschieten van computers,

hetzij

- (1) door hun beperkingen

of

- (2) doordat zij niet goed genoeg geprogrammeerd zijn – zie ook (I)

Zo te zien is er al geen sprake van gescheiden onvolkomenheden: kunstmatige en menselijke intelligentie komen met elkáár in botsing.

Waar het nu om gaat is dat deze tekortkomingen erkend worden en gepoogd wordt er 'iets aan te doen'. Zijn hier voorbeelden van te geven? Ik begin met het eerste deel van de eerste tekortkoming:

- (I) (1) mensen schieten tekort door hun onvermogen om de verrichtingen van de computer na te rekenen, terwijl daar toch fouten in kunnen zitten

Als voorbeeld kies ik het onderzoek van de wiskundige Pieter Collins en de informaticus Peter Grünwald die in september 2004 een VIDI-beurs toegekend kregen. Het is verleidelijk om het persbericht van CWI in zijn geheel te citeren, maar ik beperk me allereerst tot twee citaten:

Collins werkt aan wiskundige methodes waarmee de betrouwbaarheid van computerberekeningen bepaald kan worden. Nu is het zo dat computerberekeningen lang niet altijd nauwkeurig zijn en dat is onacceptabel als bijvoorbeeld veiligheid in het geding is. Om tot een preciezer zekerheidsmarge te komen, combineert Collins twee totaal verschillende wiskundige gebieden: de systeem- en regeltheorie die beschrijft hoe systemen reageren op signalen en de topologie, die de overeenkomst en verschillen tussen vormen beschrijft.

Grünwald gaat er in zijn onderzoek van uit dat de huidige wetenschappelijke modellen de werkelijkheid over het algemeen gebrekkig weergeven en dat de meeste modellen te complex zijn. Juist doordat ze erg volledig willen zijn en goede voorspellingen willen doen, zitten ze vol fouten en zijn ze onbetrouwbaar. Het is de kunst om de juiste balans tussen simpliciteit en complexiteit te vinden. Met behulp van statistiek, informatie-theorie en informatica wil Grünwald daartoe een algemene theorie ontwikkelen.

Goed dat erkend wordt dat computerberekeningen onnauwkeurig kunnen zijn. Mooi is het idee dat 'echte' wiskunde, met menselijke intelligentie gehanteerd, er voor zou kunnen zorgen dat fouten worden gecorrigeerd. Maar toch is het ook nog maar een idee dat nog nader uitgewerkt moet worden.

Vervolgens het tweede deel van de eerste tekortkoming:

- (I) (2) mensen schieten tekort doordat zij onvoldoende afgestemd zijn op het werken met computers en zelf fouten maken

Allereerst een simpel voorbeeld:

I have lost all my files last month. I was upgrading my computer hard disk. I forgot to back up my data. All my files gone. What a shame, although I have a Masters degree in computer science, I ignored the most basic step to safe guide your data.<sup>30</sup>

In de literatuur gaan de meest besproken voorbeelden echter over veiligheidskritische systemen. Wie kent niet de Therac-incidenten? Het is begrijpelijk als oorzaken van menselijke fouten bij toepassingen in het ‘cognitieve vlak’ liggen:

Applicatiefouten ontstaan als de verwachtingen van de gebruiker ten aanzien van het gedrag van het toestel niet in overeenstemming zijn met de werkelijkheid. Hierdoor kan de patient onnodig in zeer gevaarlijke situaties raken.<sup>31</sup>

Het lijkt er op alsof gebruikers niet op de hoogte hoeven te zijn van de mogelijkheden en onmogelijkheden van de systemen waar ze mee werken, maar de ontwerpers kunnen dat ook niet zodra het gaat om voldoende complexe systemen. Wie herinnert zich niet het ‘millenniumprobleem’? Toen wisten ontzaglijk veel mensen niet eens of hun computersysteem niet de mist in zou gaan bij de overgang van 1999 op 2000. Zolang zulke onnozelheden nog voor kunnen komen, heeft het weinig zin om optimistische verhalen over ‘kunstmatige intelligentie in de toekomst’ te verspreiden. Bovendien is er nog te weinig bekend over de interactie tussen menselijke en kunstmatige intelligentie om er voor te kunnen zorgen dat de kans op menselijke fouten zo klein mogelijk wordt. Bij veiligheidskritische apparatuur kan dit problematisch worden. Men zou willen dat mensen niet alleen op de hoogte te zijn van de mogelijkheden ervan en een morele afweging te hebben gemaakt van de mogelijke consequenties, maar ook controle kunnen blijven uitoefenen op de uitvoering.

Nu het eerste deel van de tweede tekortkoming:

(II) (1) computers schieten tekort door hun beperkingen

Dit is een stokpaardje van mij, dat gevoed werd toen ik een collega vroeg om een computerprogramma te schrijven voor een schaakcompetitieschema voor 8, 10, 12, ... spelers. Zijn eerste poging leverde een programma op dat afhaakte bij meer dan 20 spelers en een verbeterde versie met een ‘slimme’ procedure kon niet meer met 200

---

<sup>30</sup> Michael Chan, Malaysia.

<sup>31</sup> F. Koornneef en R. P. van Wijk van Brievingh.

spelers overweg. En dat terwijl de wiskundige oplossing, die met menselijke intelligentie werd gevonden, in één klap fraaie (cyclische) oplossingen voor elk even natuurlijk getal geeft.<sup>32</sup> Dit was een kwestie van beperkingen in capaciteit van computers, maar tevens kan de vraag gesteld worden of het wel mogelijk is een computerprogramma te schrijven dat die wiskundige oplossing genereert. Mijn skepsis hierover is groot ...

Tenslotte het tweede deel van de tweede tekortkoming:

- (II) (2) computers schieten tekort doordat zij niet goed genoeg geprogrammeerd zijn

Eerder heb ik met zoveel woorden opgemerkt dat dankzij de natuurlijke intelligentie van programmeurs optimale oplossingen zijn gevonden. Maar dat betref in hoofdzaak wiskundige problemen, want zodra het gaat om ‘toegepaste informatica’ blijken ‘optimale’ oplossingen van problemen ver te zoeken. En dat terwijl die ‘toegepaste informatica’ ook wel omschreven wordt als ‘informatica en softwaretechnologie ten dienste van mens en maatschappij’. Alleen al het feit dat er geen optimale spellingcontroleprogramma’s bestaan, zouden de optimisten aan het denken moeten zetten, om maar te zwijgen van complexere programma’s die door Bill Gates en consorten worden afgeleverd. Probeer maar eens een rijtje e-mailadressen in een Word-document te typen!

## Conclusies

Eerder heb ik verwezen naar een bijdrage van Baeten en Borghuis over de vraag of computers zouden kunnen rechtspreken. Welnu, een van hun argumenten voor de stelling dat computers niet zullen rechtspreken is dat ‘we de software daarvoor domweg niet kunnen produceren’, een ander, dat er ‘in het recht echter vaak sprake [is] van een niet eenduidig en soms zelfs intern inconsistent stelsel van regels’.<sup>33</sup> Van inconsistentie zou ik liever niet spreken bij rechtsregels die in gewone taal zijn geformuleerd. Wel is het zo dat zulke regels telkens geïnterpreteerd moeten worden. En dat is mensenwerk.

Anders staat het met wiskundige computerprogramma’s die gebaseerd zijn op algoritmen voor exacte domeinen, zoals ook Weizenbaum ooit moest toegeven. Er zijn misschien auteurs die dit flauw vinden, omdat het hier niet om ‘echte’ kunstmatige intelligentie (*genuine artificial intelligence*) zou gaan,<sup>34</sup> hoewel McCarthy zich hierover toch ook al genuanceerd heeft uitgelaten. Maar in het begin heb ik al aangegeven

---

<sup>32</sup> Zie mijn *Transpositions*.

<sup>33</sup> Baeten en Borghuis, o.c.

<sup>34</sup> Alex J. Champenrad..



dat ik ‘zelfs’ rekenprogramma’s al kunstmatig intelligent wil kunnen noemen. Sterker nog: aangezien de grote kracht van succesvolle computerprogramma’s gevormd wordt door algoritmen, zouden we eigenlijk dáár hun ‘kunstmatige intelligentie’ in moeten zien. Dit is in overeenstemming met de opvatting van de eerder genoemde Jan Amkreutz, die in ‘digitale algoritmen’ de principiële vernieuwing ziet van computertechnologie ten opzichte van de traditionele ‘fysische’ artefacten:

With digeality, a new type of artifact arrived: the digital machine or *thinking machine*. The electronic bank transaction or the online order that we initiate with the click of the mouse switches on a digital machine of amazing complexity. We only observe the surface – the renditions – that these machines produce; we are hardly aware of the intelligence – the algorithm – that are being executed in the digital world of bits and bytes.<sup>35</sup>

Hiermee zijn we terug bij het begin. In een ander verband heb ik mij al uitgesproken over de onwenselijkheid dat mensen de mogelijkheid verliezen om taken te vervullen die door computers kunnen worden overgenomen, maar toen ging het mij voornamelijk om cognitieve kwesties. Ik ben nu in het stadium gekomen dat ik mij zorgen maak om morele kwesties die daarmee verbonden kunnen zijn. Laten we nooit vergeten dat ‘Computers door mensen worden ontworpen en gebouwd om vervolgens programma’s uit te voeren die door mensen worden bedacht’.<sup>36</sup>

Wat hebben we tot nu toe bereikt? Ik formuleer tien stellingen:

Stelling 1. Voor het aanleren en aanwenden van automatismen is ook intelligentie vereist

Stelling 2. ‘Systematisch uitproberen of zoeken’ is ook een intelligente oplossingsmethode

Stelling 3. Mensen worden tekort gedaan als zij niet meer oefenen in vaardigheden waar computers goed in zijn

Stelling 4. Als een programma het systematische zoeken weet te bekorten met speciale procedures, is dit te danken aan de creativiteit van de programmeur

---

<sup>35</sup> J. Amkreutz, o.c., pp. 320-321.

<sup>36</sup> Naar Jan van Eyck c.s., in de prospectus van hun boek *Denkende machines*. Amsterdam: Amsterdam University Press, 2004.

Stelling 5. Alle kunstmatige intelligentie is afgeleide automatische intelligentie

Stelling 6. Systematisch zoeken is aan grenzen gebonden

Stelling 7. Problemen verzinnen is een typisch menselijke bezigheid

Stelling 8. Er zijn geen algoritmen om alledaagse problemen op te lossen

Stelling 9. Bij veiligheidskritische apparatuur is het gewenst dat mensen niet alleen op de hoogte zijn van de mogelijkheden ervan en een morele afweging hebben gemaakt van de mogelijke consequenties, maar dat zij ook controle kunnen uitoefenen op de uitvoering

Stelling 10. Ongeconditioneerd vertrouwen op computers is misplaatst.