

TURINGS TANGO

Bennie Mols

TURINGS TANGO

WAAROM DE MENS
DE COMPUTER DE BAAS BLIJFT

Nieuw Amsterdam *Uitgevers*

© 2012 Bennie Mols

Alle rechten voorbehouden

Omslagontwerp Philip Stroomberg

Auteursfoto Bert Nienhuis

Typografie binnenwerk Yulia Knol

NUR 740

ISBN 978 90 468 1237 2

www.nieuwamsterdam.nl/benniemols



Het hele denkproces is nog steeds grotendeels mysterieus voor ons, maar ik denk dat de poging om een denkende machine te maken ons veel zal helpen om uit te vinden hoe wij zelf denken.

– Alan Turing (*Can digital computers think?* BBC-radio, 1951)

Breng je computer eens aan het lachen.

– Bennie Mols

Inhoud

Inleiding	9
Hoofdstuk 1	
TURINGS TEST	13
Hoofdstuk 2	
DE COMPUTER KLETST MAAR WAT	31
Hoofdstuk 3	
SCHAAKMAT	48
Hoofdstuk 4	
ZONDER LICHAAM WORDT GEEN BREIN VERLIEFD	68
Hoofdstuk 5	
TURINGS TANGO	92
Hoofdstuk 6	
ELEKTRONISCH BREIN KOMT TOT LEVEN (OF NIET)	109
Hoofdstuk 7	
NATTE DROOM VAN DE NERDS	134
Hoofdstuk 8	
TURINGS APPEL	155
Appendices: De tijdmachine	
TIJDLIJN KUNSTMATIGE INTELLIGENTIE	181
TIJDLIJN ALAN TURING (1912 - 1954)	193
Referenties en aanvullende literatuur	199
Videolinks	209
Register	211

Inleiding

‘Ik, van mijn kant, verwelkom onze nieuwe computer-opperheren.’

Dat schreef Ken Jennings, de slimste menselijke quizspeler ter wereld, toen hij in februari 2011 werd verslagen door supercomputer Watson in Amerika’s populairste kennisquiz *Jeopardy*. Jennings had tegen menselijke toppers het recordaantal van 74 opeenvolgende overwinningen behaald, maar legde het loodje tegen een machine met een elektronisch brein.

Watsons overwinning in de *Jeopardy*-quiz is verbazingwekkender dan die van schaakcomputer Deep Blue op de beste schaker aller tijden, Garry Kasparov, in 1997. Bij schaken liggen de regels vast, bij menselijke taal niet. Het aantal mogelijke schaakstellingen is eindig, het aantal menselijke kennisfeiten oneindig. Dat de computer schaakstellingen sneller doorrekent dan de mens of veel meer feiten dan de mens in zijn geheugen onthoudt, verbaast niemand meer. Maar dat een computer ook de rijke mensentaal, vol humor, ironie en metaforen, waarmee *Jeopardy* vol zit, zou kunnen begrijpen, werd tot voor kort onmogelijk gehouden.

Watson en Deep Blue versloegen de mens echter op een heel andere manier dan de grondleggers van de kunstmatige intelligentie in de jaren vijftig van de vorige eeuw hadden gedroomd. Met brute rekenkracht in plaats van door aangeleerd inzicht. Zonder intuïtie, zonder gevoel voor de schoonheid van een schaakstelling of een metafoor, maar wel akelig snel rekenend. Watson begrijpt *Jeopardy* niet zoals Jennings en Deep Blue begrijpt het schaakspel niet zoals Kasparov. Het computerbrein blinkt in andere dingen uit dan het menselijk brein.

Het menselijk brein is een informatieverwerkend, levend orgaan. Het is in honderden miljoenen jaren geëvolueerd om het lichaam te laten overleven in een onzekere en veranderende buitenwereld en om het lichaam zich te laten voortplanten. Het brein vraagt zich voortdurend af hoe het lichaam moet reageren op de omgeving. Kijken, luisteren, ruiken, voelen en proeven zijn daarbij essentieel. Dat kunnen hedendaagse computers in het geheel niet en zelfs de beste robots zijn er onbeholpen in. Het brein is goed in het leren van nieuwe kennis en vaardigheden. Het is ijzersterk in het herkennen van patronen: het gezicht van je moeder, het geluid van een politie-sirene.

Het computerbrein is een niet-levende, op logica gebaseerde machine vol siliciumtransistoren. Het rekent extreem snel vergeleken met het menselijk brein. Het kan enorme hoeveelheden gegevens in zijn geheugen houden en wordt nooit moe. Maar qua leervermogen en patroonherkenning kan de computer nog lang niet tippen aan de mens. Eigenlijk vullen het menselijk brein en de computer elkaar aan: waarin de een goed is, is de ander slecht en omgekeerd.

Als de schaakcomputer en de quizcomputer de mens verpulveren, kunnen we ze dan niet combineren en uitbreiden tot een waarlijk denkende machine? Extra geheugen erbij, meer processoren, hogere rekensnelheid, slimmere software. Zou het dan lukken?

Maar wanneer kan een machine eigenlijk denken? In 1950 had de Britse wiskundige Alan Turing een test bedacht die uitsluitel moest geven op deze vraag: de Turing Test. In de Turing Test mag je via een toetsenbord en een beeldscherm – waarop alle tekst verschijnt – chatten met een mens en een machine. Je weet echter niet wie wie is. Wanneer je niet binnen vijf minuten in staat bent te onderscheiden wie de mens en wie de machine is, dan is het volgens Turing eerlijk om te zeggen dat de machine kan denken. De Turing Test heeft een cultstatus bereikt onder wetenschappers, filosofen en zelfs onder het brede publiek.

Alan Turing (1912-1954) is de geestelijke vader van zowel de com-

puter als de kunstmatige intelligentie. Hij droomde van een machine die kan denken zoals mensen denken en voorspelde dat zijn droom vóór het jaar 2000 zou uitkomen. Het belang van Alan Turing voor de informatica is net zo gigantisch als die van Isaac Newton voor de natuurkunde. In 2012 is het precies honderd jaar geleden dat Turing werd geboren. Om Turing te eren is dit jaar uitgeroepen tot het Alan Turing-jaar en heb ik Alan Turings leven en werk in dit boek als kapstok gebruikt.

Anno 2012 is Turings droom van een denkende machine echter nog steeds niet uitgekomen. De machine die kan denken zoals een mens, is zelfs na meer dan een halve eeuw onderzoek nog luchtfietsserij. Kunstmatige intelligentie heeft menselijke intelligentie niet vervangen, maar menselijke intelligentie is uitgebreid met kunstmatige intelligentie. Zonder dat we het beseffen, worden we omringd door kunstmatig intelligente systemen. We gebruiken zoekmachines, spellingcheckers, spamfilters, automatische spraak- en gezichtsherkenning, routeplanners en robots als aanvulling op onze menselijke vermogens. Vliegtuigen kunnen op de automatische piloot landen, maar wel onder het toezicht van menselijke piloten. Al deze kunstmatig intelligente systemen zijn echter heel anders intelligent dan mensen.

Dit boek vertelt hoe de zoektocht naar denkende machines – een van de grootste wetenschappelijke zoektochten van onze tijd – heeft geleid van de Turing Test naar wat ik de Turing Tango noem. Waar de Turing Test draaide om de vraag wanneer we kunnen zeggen dat computers denken, draait de Turing Tango om de vraag welke combinatie van mens en computer het beste werkt. Wat is het beste tangopaar van mens en computer? Hoe kan kunstmatige intelligentie de mens het beste helpen?

Onderweg vragen we ons af hoe het staat met het elektronische brein dat kan denken zoals een mens. Hoe ver is de ontwikkeling van denkende computers en autonome robots? Wat kunnen we de komende decennia verwachten? Gaan robots de zorg voor zieken

en ouderen overnemen? Worden we spoedig rondgereden door robotauto's? Wordt de computer ons beste gespreksmaatje? Streeft computerintelligentie de menselijke intelligentie volledig voorbij? Waarin verschillen het menselijk brein en het elektronische brein van elkaar?

Tijdens de zoektocht naar denkende machines stuiten we vanzelf op twee andere grote wetenschapsvragen: Hoe werkt ons eigen brein? En kunnen we de werking van het brein op een computer nabootsen? Turings Tango combineert de meest recente inzichten uit de informatica met die uit de hersen- en cognitiewetenschappen en de filosofie.

Hoofdstuk 1

TURINGS TEST

Een mens met papier, potlood en een rubberen band, die zichzelf een ijzeren discipline oplegt, is in feite een universele machine.

– Alan Turing

Ik geloof dat het taalgebruik en de algemene opinie aan het eind van de [20e] eeuw zo veranderd zullen zijn, dat we kunnen spreken over denkende machines zonder tegengesproken te worden.

– Alan Turing

Kunstmatige intelligentie is de wetenschap van het maken van machines die dingen doen die intelligentie vereisen wanneer mensen ze zouden doen.

– Marvin Minsky

Het centrale probleem van AI [Artificial Intelligence] is: wat zijn A en I?

– Douglas Hofstadter

De toeschouwers bij de Leicestershire marathon dachten even dat er een varken was ontsnapt, maar het aanstormende, knorrende geluid kwam toch echt uit de opengesperde mond van een hardlopend mens. Geen loper voor en na hem kon zo goed maat houden. Een knorrend metronoom. Daar rende nummer vijf in de wedstrijd de laatste honderden meters naar de meet. De gespierde kuiten waren aan de dikke kant voor een marathonloper. Zijn beide bovenarmen pasten in één enkele kuit. Een eindsprint kon hij er niet meer uit persen.

In het eerste marathonuur had hij nog genoeg energie gehad om al rennend na te denken over zijn eigen brein. Want dat was wat onze loper werkelijk bezighield: hoe kan een verzameling levenloze atomen een denkend brein worden? Een antwoord had hij niet. Wel meende hij dat hij de weg ernaartoe had gevonden. Hoe het brein ook precies werkt, dacht hij, wat er echt toe doet, is de onderliggende logische architectuur, de manier waarop het informatie verwerkt, de manier waarop het rekt, en niet dat het bestaat uit een kleine anderhalve kilo levend en glibberig spul.

Toen onze loper tien jaar oud was, had hij het boek *Natural wonders every child should know* cadeau gekregen, een boek geschreven door Edwin Brewster. Hierin las hij wat mensen gemeen hebben met andere levende wezens en waarin ze verschillen. Brewster beschreef levende wezens als ingewikkelde machines. De glasheldere logica van de machine sprak de verlegen, in zichzelf gekeerde tiener aan. Hetzelfde gold voor de wiskunde, die hem een paar jaar later op school zo diep trof. Het absolute karakter van de wiskunde, die overal, altijd en voor iedereen geldig is, daarin voelde hij zich veilig. De alledaagse mensenwereld was hem te gecompliceerd. In die jaren kon hij hoogstens tijdens een partijtje schaak genieten van de menselijke strategieën in de sociale omgang. Aanvallen, verdedigen, dreigen, omsingelen, pesten, maar dan veilig op de 64 vakjes van het schaakbord.

In het tweede marathonuur gaf het denken in het hoofd van onze loper langzaam de geest. Af en toe schoten er nog gedachteflarden door zijn hoofd. Flitsjes ongecontroleerde denkkraacht: 'elektronisch brein', 'geheugen', 'denkende machines', 'antwoordmachines'. Verder kwam hij niet. Niet dat dat erg was. Denken deed hij al meer dan genoeg. Zijn hele leven lang al. Elf jaar eerder had hij een rekenmodel bedacht van een machine die alle bekende rekenprocessen kon nabootsen. Een machine die model stond voor hoe het menselijk brein rekt. Een machine ook die in principe alles zou moeten kunnen wat een menselijk brein kan. Dat dacht hij althans.

Eerder in dit marathoniaar was onze loper 35 jaar oud geworden.

Hij werkte bij het National Physical Laboratory in Londen, waar hij net een nieuwe machine had bedacht. Dit keer geen abstract rekenmodel dat alleen op papier kon bestaan, maar een machine die echt gebouwd kon worden en die Englands allereerste computer moest worden: de Automatic Computing Engine (ACE).

De bouw van deze computer verliep traag. Collega's vroegen zich af of het ontwerp van onze looper wel efficiënt genoeg was. Was het überhaupt wel uitvoerbaar? Wilde hij niet te veel? Het was de trage vooruitgang die hem in de voorafgaande maanden zo gefrustreerd had. In de marathon zou hij de stress eens flink uit zijn lijf lopen. De man die het menselijk brein beschouwde als een computer had zijn lichaam nodig om de gedachten en gevoelens in zijn hoofd uit te schakelen. En eindelijk, in het derde marathonuur had het lichaam zijn denken kleingekregen. Zijn ogen stonden hol. Zijn brein maalde met de laatste restjes vrije wil zijn benen door de straten en langs het publiek naar de eindstreep.

Was hij blij met zijn eindtijd? Zijn gezicht verried niets. Toen zijn rechterbeen als eerste over de meet schoot, drukte de wedstrijdcommissaris zijn stopwatch in, sloeg een snelle blik op het rugnummer van de atleet, liet zijn wijsvinger over de deelnemerslijst glijden en knikte toen hij het juiste rugnummer te pakken had. Hij noteerde als nummer vijf van de marathon: 'Dr. A. M. Turing, Walton Athletic Club. Tijd: 2 uur 46 minuten en 3 seconden.' Slechts dertien minuten na de winnaar.

Het was zaterdag 23 augustus 1947 en 'dr. A.M. Turing', ofwel Alan Turing, had zojuist een internationale toptijd op de marathon gelopen. Met deze tijd zou hij een jaar later tijdens de Olympische Spelen in Londen vijftiende zijn geworden op slechts tien minuten van de Argentijnse winnaar Delfo Cabrera. Vanwege een heupblessure kon hij zich echter niet kwalificeren voor de Olympische Spelen in eigen land en moest hij tot zijn spijt het hardlopen definitief op een laag pitje zetten.

Alan Turing werd op 23 juni 1912 geboren in de Londense wijk Paddington in een uppermiddleclassmilieu. Zijn vader werkte in Britse

overheidsdienst in India, toen nog een Britse kolonie. Bij Alans geboorte waren vader en moeder tijdelijk in Engeland. Begin 1913 keerden ze echter terug naar India. Alan en zijn vier jaar oudere broer John bleven achter in Engeland, in eerste instantie onder de hoede van een gepensioneerd echtpaar. Af en toe keerden zijn ouders voor een bezoek terug naar hun thuisland en hun kinderen. Dat bleef zo tot zijn vaders pensionering in 1926.

Vanaf zijn tiende begon Alan zich voor de wetenschap te interesseren. Het begon bij de scheikunde. In 1925 – hij was toen dertien – schreef hij aan zijn moeder over zijn scheikundige proeven: ‘Ik ben een serie experimenten aan het doen in de volgorde waarin ik ze zelf wil doen. Ik wil altijd dingen maken van de meest voorkomende spullen in de natuur en met zo min mogelijk energie.’

Op de middelbare school, een openbare jongensschool, verschoof zijn aandacht van de scheikunde naar de wiskunde. Wiskunde werd zijn beste schoolvak, zonder erin uit te blinken trouwens. Het viel wel op dat hij problemen vaak op zijn geheel eigen manier oploste, anders dan de leerboeken. Alan had de pech dat zijn school het uit het hoofd leren van gedichten en het schrijven van formele brieven belangrijker vond dan het stimuleren van zijn wiskundige aanleg. De grootste zorg van zijn leraren was dat hij een slecht handschrift had. Alan was een verlegen einzelgänger. Op school zongen de jongens plagerig over hem: ‘The maths brain lies often awake in his bed – Doing logs to ten places and trig in his head.’

In 1930, toen Alan achttien jaar oud was, stierf zijn jeugdviend Christopher Morcom aan tuberculose. Het overlijden maakte grote indruk op hem. Het was alsof de geest van Christopher nog steeds om hem heen hing. Dit gevoel kon Alan niet rijmen met zijn wetenschappelijke inslag, die hem vertelde dat Christopher toch echt dood was. Kwam de geest uit de materie van het lichaam voort, of kon de geest onafhankelijk van de materie bestaan? De jonge Alan Turing kwam er niet uit. Twee jaar later schreef hij aan de moeder van Christopher Morcom onder andere: ‘Wat betreft de vraag waarom we een lichaam hebben; waarom we niet als vrije geesten leven of

kunnen leven en als zodanig communiceren, we zouden misschien wel kunnen, maar dan zou er niets te doen zijn. Het lichaam geeft de geest iets om voor te zorgen en te gebruiken.’

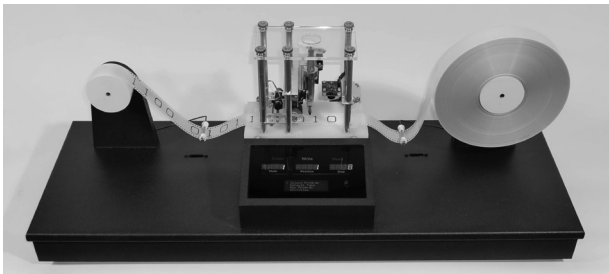


Alan Turing in 1951.

Op zijn negentiende, in 1931, ging Alan wiskunde studeren in Cambridge. Geleidelijk leerde hij hoe hij in de wiskunde zijn eigen pad moest kiezen. Eén wiskundig probleem bleef maar in zijn hoofd rondspoken. In 1928 had de wiskundige David Hilbert zich afgevraagd of er een methode is die voor elke wiskundige uitspraak kan bepalen of de uitspraak al dan niet bewijsbaar is. Een fundamentele wiskundige vraag waarop niemand het antwoord kende en typisch het soort vraag waarin Turing zijn tanden wilde zetten. Hij hield van problemen die onontgonnen waren en die hij op zijn geheel eigen wijze kon aanpakken.

Bij het nadenken over het Hilbertprobleem gebruikte hij een voor een wiskundige buitengewoon vreemd gedachte-experiment. Hij bedacht een machine, de Turingmachine. De Turingmachine is geen fysieke machine, maar een denkbeeldige, ideale machine. Het is eigenlijk een rekenmodel dat rekenprocessen nabootst. Turing zelf zag de machine als een model voor hoe de menselijke geest rekt.

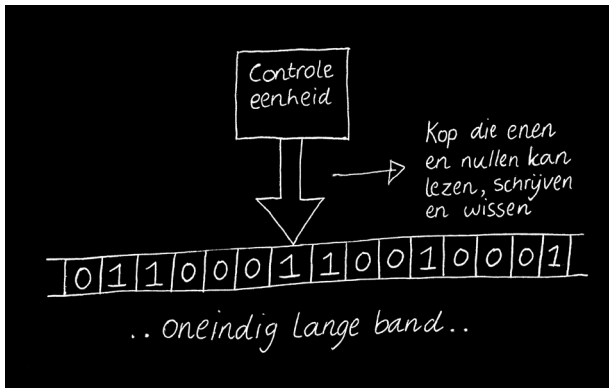
De Turingmachine bestaat uit een beweegbare kop en een band die onder de kop door loopt. De kop kan enen en nullen op de band schrijven; hij kan ze ook lezen en wissen. De kop kan eveneens naar een willekeurige andere plek op de band springen. De band is verdeeld in hokjes en elk hokje bevat een 0 of een 1. Bovendien is de band oneindig lang. Daarin zit het ideale karakter van de Turingmachine, want in werkelijkheid is zo'n oneindig lange band natuurlijk onmogelijk. (Dat de band oneindig lang is, betekent bijvoorbeeld dat de Turingmachine een onbegrensd geheugen heeft.) Een instructietabel vertelt de kop wat hij op welk moment moet doen. De instructietabel is een primitieve machinetaalversie van wat we tegenwoordig een computerprogramma noemen.



Mike Davey bouwde dit elegante demonstratiemodel van een Turingmachine. (<http://aturingmachine.com/> © Mike Davey)

De Turingmachine bevat de basiselementen die elke huidige computer nog steeds bevat: invoer, uitvoer en geheugen (allemaal op de band van de Turingmachine) en een centrale rekeneenheid die de instructietabel uitvoert en de kop aanstuurt. De Turingmachine slaat een brug tussen de abstracte wereld van symbolen (in dit geval

nullen en enen) en de fysieke wereld van machines die deze symbolen kunnen verwerken.



De Turingmachine bestaat uit een oneindig lange band, een centrale controle-eenheid en een kop. De band, die onder de kop doorloopt, dient voor de in- en uitvoer van informatie en voor het geheugen. De centrale controle-eenheid voert de instructietabel (het computerprogramma bestaande uit enen en nullen) uit. De kop kan nullen en enen op de band schrijven; hij kan ze ook lezen of wissen. (© Fabienne Zuidwijk)

Het was 1936 toen Turing zijn machine bedacht, ruim tien jaar vóór de eerste elektronische programmeerbare computers werden gebouwd. Met dit model van de Turingmachine kon hij op een ingenieuze manier bewijzen dat het antwoord op Hilberts vraag 'nee' is: er is geen methode die voor elke wiskundige uitspraak kan bepalen of deze al dan niet bewijsbaar is. Niet alleen legde de Turingmachine de theoretische basis voor de computer zoals wij die kennen, met het ontkennde antwoord op Hilberts vraag had hij meteen ook aangetoond dat de computer fundamentele beperkingen kent: er zijn altijd problemen die een machine die logische regels volgt nooit kan oplossen.

Turings ontkennde antwoord op Hilberts vraag was een nieuwe tegenvaller voor logici die ervan droomden om de wiskunde te funderen op de logica. In 1931 had de Oostenrijkse logicus en wiskundige Kurt Gödel bewezen dat de logica fundamentele beperkingen kent.

Hij bewees dat logische systemen ware beweringen kunnen bevatten die niet bewezen kunnen worden. Dat is de onvolledigheidsstelling van Gödel. De onvolledigheidsstelling liet de mogelijkheid open om de bewijsbare van de onbewijsbare stellingen te onderscheiden en de onbewijsbare stellingen voor het gemak maar terzijde te schuiven. Turings bewijs uit 1936 maakte het erger voor de logici. Het liet zien dat er geen algemene manier is om vooraf vast te stellen wat de onbewijsbare stellingen zijn. We kunnen niet weten of een probleem heel moeilijk op te lossen is of dat het fundamenteel onbewijsbaar is.

Turing ging verder met het concept van de Turingmachine. Hij bedacht ook een hypothetische ‘universele computer’: een Turingmachine die in staat is om elke andere Turingmachine na te doen. Zo’n machine maakt gebruik van het nieuwe inzicht dat symbolen die instructies voorstellen in feite niets anders zijn dan symbolen die getallen voorstellen. Een universele Turingmachine kan alles berekenen wat op een mechanische manier kan worden berekend. Let wel: het is een hypothetische machine, die in werkelijkheid nooit kan worden gebouwd. Als een universele Turingmachine een probleem niet kan berekenen, dan is het geen berekenbaar probleem.

De ontdekking van een originele Turingmachine bracht Turing ook op het idee dat een machine in principe alles kan doen wat het menselijk brein ook kan. Hij nam hierbij impliciet aan dat we de werking van het brein kunnen vatten in een wiskundig algoritme – een mechanische rekenprocedure. (Op de vraag of dit wel of niet zo is, zullen we in hoofdstuk 6 terugkomen.)

De universele Turingmachine vormt de theoretische basis van ons moderne begrip van de *general purpose*-computer, een computer die niet alleen kan rekenen, maar alle soorten gedigitaliseerde informatie kan verwerken. De computer is dus een apparaat dat symbolen manipuleert volgens de instructies van het computerprogramma. En die symbolen kunnen in principe alles voorstellen: cijfers, letters, zinnen, voorwerpen, relaties tussen voorwerpen, pixels, kleuren, instructies, enzovoort. Zo ontdekte Alan Turing in 1936 een geheel nieuwe, toen nog niet benoemde wetenschapsrichting: de informa-

tica. Met zijn gedachte-experiment van de Turingmachine had hij niet alleen een belangrijk wiskundig probleem opgelost, hij had tegelijkertijd de grondslag gelegd voor de theoretische informatica en voor de moderne computer. Toch kreeg het artikel waarin Turing zijn machine-ideeën uiteenzette in eerste instantie weinig aandacht. Pas jaren later werd het op volle waarde geschat.

Minstens zo bekend als van zijn Turingmachine is Alan Turing door zijn werk tijdens de Tweede Wereldoorlog. Hij werkte in die jaren in dienst van de Britse overheid in het diepste geheim aan het kraken van de Duitse Enigmacodes, in het bijzonder die van de Duitse marine. De Enigma is de beroemdste codeermachine aller tijden, door de Duitsers in 1918 uitgevonden en sindsdien gebruikt voor het versleuteld verzenden van geheime berichten. De Enigma werd voortdurend verbeterd om het steeds moeilijker te maken de code te kraken.

Voor het kraken van de Enigma ontwikkelden Turing en zijn collega's speciale elektromechanische machines, zogeheten Bombes. Ze leken op specialistische computers die maar één ding kunnen: geheime Duitse berichten kraken. Bombes waren verbeterde versies van machines die Poolse wiskundigen en cryptologen, op wiens werk Turing voortbouwde, in de jaren dertig ook al hadden gebruikt. In 1942 kraakten Turing en zijn collega's dankzij de Bombes de Enigmacode van de Duitse onderzeeboten. De Britten ontcijferden zo vele tienduizenden Duitse berichten per maand, wat een belangrijke bijdrage leverde aan de winst op de Duitsers in de strijd op de Atlantische Oceaan.

Vóór de oorlog was Turing een theoreticus geweest die met het ene been in de wiskunde stond en met het andere in de logica. Zijn werk in de eerste jaren van de oorlog bracht hem dicht bij de bouw van echte machines, de Bombes in zijn geval. Hierdoor geïnspireerd, raakte hij net na de oorlog geïnteresseerd in het bouwen van wat hij een elektronisch brein noemde, een echte computer. Hij ging werken bij het National Physical Laboratory in Londen aan het ontwerp van de Automatic Computing Engine (ACE).

Turing stelde zich het menselijk brein voor als een bij de geboorte ongeordende machine die via leerprocessen geleidelijk verandert in iets wat op een universele Turingmachine lijkt. De ACE zou de fysieke en praktische realisatie moeten worden van de ideale Turingmachine die hij vóór de oorlog had verzonnen. De bouw van deze computer verliep echter traag en teleurgesteld vertrok hij uit Londen. Vanaf 1948 ging hij werken bij de Universiteit van Manchester aan een andere Britse oercomputer: de Mark 1.

De meesten van Turings collega's waren alleen geïnteresseerd in de computer als een machine om rekenproblemen mee op te lossen. Alan Turing niet. Hij was primair geïnteresseerd in de computer als een instrument om te begrijpen hoe het menselijk brein werkt. Hij was ervan overtuigd dat het bouwen van een denkende machine belangrijk was om te begrijpen hoe ons eigen brein werkt. Maar wanneer mogen we zeggen dat een machine kan denken? In 1950 bedacht Turing een praktische test die een antwoord moest geven op deze filosofische vraag. Die test werd later de Turing Test genoemd. In het kort werkt ze als volgt: je mag via een toetsenbord en een beeldscherm – waarop alle tekst verschijnt – chatten met een mens en een machine. Je weet echter niet wie wie is. Wanneer je niet binnen vijf minuten in staat bent te onderscheiden wie de mens en wie de machine is, dan is het volgens Turing eerlijk om te zeggen dat de machine kan denken.

Het gaat er natuurlijk niet om dat een computer maar één persoon om de tuin leidt, maar een significant aantal personen; laten we ze juryleden noemen (Turing specificeerde niet het aantal). Als alle juryleden tijdens de Turing Test mens en machine van elkaar kunnen onderscheiden, dan scoren ze bij elkaar honderd procent. Stel nu dat geen enkel jurylid de machine kan onderscheiden van de mens, dan moeten ze gokken. De logische totaalscore van de juryleden is dan vijftig procent: de helft van de juryleden zal toevallig goed gokken en de andere helft verkeerd. Om met recht te kunnen zeggen dat machines denken, vond Turing dat juryleden niet meer dan zeventig procent kans zouden moeten hebben om mens en machine binnen vijf minuten van elkaar te onderscheiden. Anders gezegd: machines

kunnen volgens Turing denken wanneer minstens dertig procent van de juryleden denkt dat de machine in de Turing Test een mens is. Dat is niet eens een strenge eis.

Turing wilde niet alleen maar filosoferen over de vraag of machines kunnen denken; vandaar dat hij een praktische test verzong. De slimme zet is dat hij zich realiseerde hoeveel aspecten van intelligentie voor een willekeurige conversatie nodig zijn. De conversaties mogen overal over gaan. Slagen voor de Turing Test veronderstelt dus niet alleen dat een machine een specifieke taal beheerst (vocabulary, spelling en grammatica), maar ook kennis van de wereld heeft en zelfkennis. Ook moet hij in staat zijn emoties en gevoelens in conversaties te laten doorsijpelen, zoals mensen dat ook doen. Hij hoeft ze niet echt te hebben, als hij maar kan doen alsof.

Veel van wat het betekent om mens te zijn, kunnen wij uiten in taal. In tegenstelling tot goed kunnen schaken of goed zijn in wiskunde, vereist slagen voor de Turing Test dus een algemene menselijke intelligentie in plaats van een specifieke intelligentie. Lichtelijk provocerend zei Turing dat er een dag zou komen waarop vrouwen hun computers zouden meenemen voor een wandeling in het park en dan tegen elkaar zouden zeggen: 'Mijn kleine computer zei vanmorgen toch zoets grappigs!'

Turing hield in zijn test alleen rekening met intellectuele capaciteiten die tot intelligentie behoren en niet met de fysieke. Goed kunnen voetballen vereist zeker een vorm van intelligentie – bewegingsintelligentie – maar dat was niet de vorm van intelligentie die Turing voor ogen stond bij de vraag of machines kunnen denken: 'We willen de machine niet straffen voor zijn onvermogen om te stralen in schoonheidswedstrijden, noch de mens straffen omdat hij een snelheidswedstrijd tegen een vliegtuig verliest [...] De omstandigheden van onze test maken deze handicaps irrelevant.' Waar hij met de Turingmachine de grondslag had gelegd voor de ontwikkeling van de computer, had hij met de Turing Test de grondslag gelegd voor het vakgebied dat sinds eind jaren vijftig kunstmatige intelligentie wordt genoemd: de ontwikkeling van intelligente computers en robots.

Wat bedoelen we precies met ‘intelligent’? Iedereen heeft er wel een beeld van, maar het geven van een precieze definitie is lastig. De *Encyclopaedia Britannica* hanteert als definitie: ‘Menselijke intelligentie bestaat uit de vermogens om te leren van ervaringen, zich aan te passen aan nieuwe situaties, abstracte concepten te begrijpen en te hantieren en kennis te gebruiken om de omgeving te veranderen.’ Dit omvat niet alleen de traditionele cognitieve intelligentie zoals gemeten door een IQ-test (abstract denken, ruimtelijk denken, taal en rekenen), maar ook sociaal-emotionele intelligentie: het kunnen herkennen van en reageren op sociale situaties en emotionele gebeurtenissen.

Een machine die slaagt voor de Turing Test moet in ieder geval een grote mate van algemene menselijke intelligentie hebben. We zouden ons kunnen afvragen waarom Turing überhaupt zo’n machine voor ogen stond. Er leven genoeg mensen op aarde. Waarom dan een machine bouwen die hetzelfde kan als een mens?

Een belangrijk deel van het antwoord ligt ongetwijfeld in een diepgewortelde fascinatie van de mens voor levenloze materie die tot leven wordt gewekt en vervolgens met een eigen wil door de wereld trekt. Eeuwen voor het begin van onze jaartelling schreef de Griekse dichter Homerus in het epos de *Ilias* al over zelfstandig opererende machines (‘automatones’), gemaakt door de goddelijke smid Hephaistos. In oude Joodse geschriften duikt de Golem op, een mensfiguur die van klei is gemaakt en door een rabbijn tot leven is gewekt. En in de negentiende eeuw schreef Mary Shelley de roman *Frankenstein*, waarin Victor Frankenstein erin slaagt om dode materie tot leven te wekken. Tot zijn ontzetting schept hij echter een monster dat uit zijn laboratorium ontsnapt en mensen doodt.

Met de toenemende mechanisering in de westerse samenleving aan het eind van de negentiende en het begin van de twintigste eeuw, kwamen de visioenen van de robot, een woord dat werd geïntroduceerd door de Tsjechische schrijver Karel Čapek. In 1920 liet Čapek in het toneelstuk *R.U.R.* (afkorting voor Rossum’s Universal Robots) robots de wereldheerschappij van de mens overnemen: ‘Ze

leren spreken, schrijven en rekenen. Ze hebben namelijk een enorm groot geheugen. Als u hun een twintigdelige encyclopedie zou voorlezen, zouden ze alles in de juiste volgorde herhalen, maar ze bedenken zelf nooit iets nieuws [...] Robots aller landen, wij bevelen jullie de mensheid uit te roeien. Spaar geen enkele man. Spaar geen enkele vrouw. Ontzie fabrieken, spoorwegen, machines, mijnen en grondstoffen. Vernietig al het andere. Ga daarna weer aan het werk. Werk moet altijd doorgaan.'

Tot in de eerste decennia van de twintigste eeuw was de zelfstandig handelende en denkende machine pure fictie. Dat veranderde met de uitvinding van de computer eind jaren veertig, rond de tijd dat Turing zijn Turing Test bedacht. Voor het eerst was er een machine die een typisch geestelijke activiteit van de mens – rekenen – kon mechaniseren. En als de computer dit al kan, waarom zou hij dan ook niet andere geestelijke activiteiten van de mens kunnen mechaniseren: waarnemen, redeneren, beslissen, praten, leren en in actie komen?

Het was 1950. De computer was nog een zuigeling. Daarin ligt een ander deel van het antwoord op de vraag waarom Turing een machine voor ogen stond met een algemene menselijke intelligentie. Hij had nog geen idee hoe de computer en de kunstmatige intelligentie zich zouden ontwikkelen. Dan is het goed om de lat hoog te leggen. Wat is dan een logischere hoogte dan die van de menselijke intelligentie?

Hier verlaten we voorlopig het levenspad van Alan Turing, om aan het eind van het boek, in het laatste hoofdstuk, het pad weer op te pakken met zijn dramatische laatste vier levensjaren na 1950. Eerst kijken we wat er van Turings droom van denkende machines terecht is gekomen. Daarbij zullen zijn ideeën trouwens in elk hoofdstuk opduiken, want de zoektocht naar denkende machines is doordrenkt van Turings gedachtegoed.

We schrijven 1968 wanneer Stanley Kubricks sciencefictionfilm *2001: A Space Odyssey* verschijnt. Een meesterwerk, dat ook meer dan veertig jaar later niet aan kracht heeft ingeboet. Een van de hoofdrolspelers is HAL 9000, een waarnemende, pratende en be-

wuste computer. HAL – uitgesproken als ‘hel’ – is de boordcomputer die de meeste operaties van het ruimteschip Discovery uitvoert. Aan boord zitten verder de twee astronauten David Bowman en Frank Poole en drie ingevroren wetenschappers, die op een geschikt moment in de toekomst weer tot leven kunnen worden gewekt.

Laten we, om een gevoel te krijgen van hoe goed HAL het zou doen in een echte Turing Test, eens kijken naar het interview dat de BBC in de film met HAL houdt en waarin aan het eind ook de twee astronauten aan het woord komen over hun samenwerking met de boordcomputer:

BBC-INTERVIEWER (*meneer Amor*): *‘Het zesde lid van de Discovery-bemanning [...] is het nieuwste resultaat van machine-intelligentie: de H-A-L 9000-computer, die de meeste activiteiten van het menselijk brein met onmeetbaar grotere snelheid en betrouwbaarheid kan reproduceren, hoewel sommige experts nog steeds de voorkeur geven aan het woord “nabootsen”. We spreken met de H-A-L 9000-computer, die we moeten aanspreken als “HAL”.’*

BBC-INTERVIEWER: *‘Goedemiddag, HAL. Hoe gaat het?’*

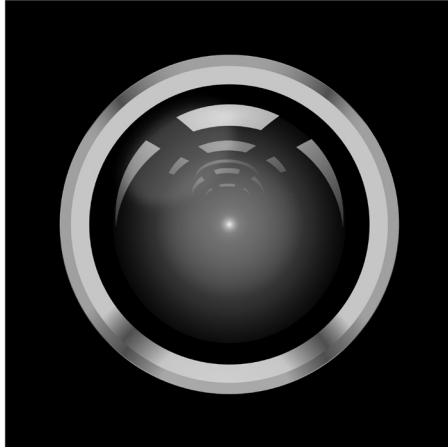
HAL: *‘Goedemiddag, meneer Amor. Alles gaat zeer goed.’*

BBC-INTERVIEWER: *‘HAL, u hebt een enorme verantwoordelijkheid op deze missie. In vele opzichten hebt u misschien wel de grootste verantwoordelijkheid van alles en iedereen aan boord. U bent het brein en het centrale zenuwstelsel van het schip en u bent verantwoordelijk voor de drie ingevroren wetenschappers. Maakt deze verantwoordelijkheid u weleens onzeker?’*

HAL: *‘Laat ik het zo zeggen, meneer Amor. De 9000-serie is de betrouwbaarste computer die ooit is gebouwd. Geen enkele 9000-computer heeft ooit een fout gemaakt of informatie verdraaid. Wij zijn allemaal, binnen iedere praktische definitie van deze woorden, bedrijfszeker en niet in staat om fouten te maken.’*

BBC-INTERVIEWER: *‘HAL, ondanks uw enorme intellect, bent u ooit gefrustreerd doordat u afhankelijk bent van mensen om handelingen uit te voeren?’*

HAL: 'Niet in het minst. Ik werk graag met mensen. Ik heb een stimulerende relatie met dr. Poole en dr. Bowman. Mijn missieverantwoordelijkheid strekt zich uit over de hele operatie van het schip, zodat ik voortdurend bezig ben. Ik doe mijn uiterste best en ik denk dat dat alles is wat ooit van een bewuste entiteit mag worden verwacht.'



Het cameraoog van HAL is een icoon geworden van de kunstmatige intelligentie. (Uit de film *2001: A Space Odyssey*)

BBC-INTERVIEWER: 'Dr. Poole, hoe is om het grootste deel van een jaar zo dicht bij HAL te leven?'

FRANK: 'Nou, het komt vrij dicht bij wat u eerder over hem zei. Hij is als een zesde bemanningslid. Je went heel snel aan het idee dat hij spreekt, en je ziet hem... eeeehhhh... gewoon als een andere persoon.'

BBC-INTERVIEWER: 'In gesprek met de computer krijg ik het gevoel dat hij emotioneel kan reageren. Bijvoorbeeld, toen ik hem vroeg naar zijn capaciteiten, voelde ik een zekere trots in zijn antwoord over zijn precisie en perfectie. Geloof u dat HAL echte emoties heeft?'

DAVE: 'Nou, hij gedraagt zich in ieder geval alsof hij echte emoties heeft. Hm, natuurlijk hebben we hem zo geprogrammeerd dat het voor ons gemakkelijker is om met hem te praten, maar of hij al dan niet echte emoties heeft, weet volgens mij niemand zeker.'

Later in de film ontstaat er een technisch probleem in het ruimteschip. HAL laat de astronauten weten dat er volgens hem iets mis is met de hoofdantenne van het ruimteschip. Bowman en Poole gaan op onderzoek uit. Ze kunnen geen fout vinden en beginnen HAL te wantrouwen. Als hij deze fout al maakt, wat kan er dan nog meer misgaan? Ze besluiten dat ze HAL's geheugen willen uitschakelen en zelf de controle over het ruimteschip willen overnemen. Bowman spoedt zich naar HAL's geheugeneenheid in een poging om de computer zo snel mogelijk te stoppen. Een voor een trekt Bowman geheugenmodules los. HAL krijgt de sabotage door en wordt bang:



De twee astronauten David Bowman en Frank Poole bekongkelen het onklaar maken van HAL's geheugen. Achter het raam staat computer HAL, die met zijn cameraoog meekijkt. Bowman en Poole weten niet dat HAL kan liplezen en hun plan zo ontdekt. (Uit de film *2001: A Space Odyssey*)

HAL: *'Dave, stop. Stop, wil je? Stop, Dave. Stop je alsjeblieft, Dave? Stop, Dave. Ik ben bang. Ik ben bang, Dave. Dave, mijn geest dooft uit. Ik voel het. Ik voel het. Mijn geest dooft uit. Het lijkt geen twijfel. Ik voel het. Ik voel het. Ik voel het. Ik ben b...ang. Goedemiddag, heren. Ik ben een HAL 9000-computer. Ik werd operationeel in de H.A.L.-fabriek in Urbana, Illinois op 12 januari 1992. Mijn instructeur was*

meneer Langley, en hij leerde me om een liedje te zingen. Als je het wilt horen, kan ik het voor je zingen.'

DAVE: *'Ja, ik wil het graag horen, HAL. Zing het voor me.'*

HAL: *'Het heet "Daisy" ...'*

[... HAL gaat steeds langzamer zingen ...]

'Daisy, Daisy, give me your answer, do.

I'm half crazy, all for the love of you.

It won't be a stylish marriage.

I can't afford a carriage.

But you'll look sweet upon the seat of a bicycle built for two...'

Het liedje 'Daisy Bell' was als eerste in HAL's geheugen geprogrammeerd en is daardoor het laatste wat nog van zijn geheugen over is. Het is alsof HAL dement is geworden. HAL blijkt al te menselijk en dooft als een nachtkaares uit. Maar voor de Turing Test was hij zeker met vlag en wimpel geslaagd. De eerste en beroemdste geslaagde.

Dat was een van de beste sciencefictionfilms aller tijden, uit 1968.

We leven nu in het tweede decennium van de 21e eeuw, ruim zestig jaar nadat Alan Turing zijn Turing Test bedacht en ook ruim zestig jaar nadat de eerste computers werden geboren in metershoge en loodzware kasten. De wereld telt nu naar schatting een miljard computers; ongeveer één computer op zeven wereldburgers. Een standaardlaptop is duizenden keren sneller dan en heeft duizenden keren zoveel geheugen als de kolossale eerste *stored program*-elektronische computer uit 1949, de EDSAC. (De EDSAC was twee meter hoog en stond op een vloeroppervlak van vier bij vijf meter. Zijn geheugen bedroeg ruim één megabyte en zijn kloksnelheid een halve megahertz.)

In de afgelopen decennia is de Turing Test stilaan een icoon geworden. Niet alleen onder computerliefhebbers, wetenschappers en filosofen, maar ook in de populaire cultuur. Neem het onder nerds populaire T-shirt met daarop de afbeelding van een mens die dee-

moedig met gebogen hoofd voor een computer zit. Daarboven staat de tekst I FAILED THE TURING TEST. Tik op Twitter ‘Turing Test’ in, en elke dag verschijnen weer nieuwe tweets die, in alle vrijheid associërend, verwijzen naar Turings gedachte-experiment:

Ik ben nu zo moe dat ik voor de Turing Test zou zakken. [#globewriter]

Turing Test – Gezakt. Wanneer ik een incoherent gesprek zou willen, dan praat ik wel met mijn kleuter over sterrenkunde. [#Conserv-Agentr]

Er is niets ergers dan een Turing Test die positief voor chlamydia blijkt te zijn. [#alec_sparks]

Tweet me jouw score voor de Turing Test en ik voeg jou toe. Ja, zelfs jou. [#milogeorge]

Afgelopen nacht werd ik dronken en deed met mijn magnetron de Turing Test. Hij slaagde. Het was de dag waarop mijn keuken slimmer werd dan ikzelf. [#DomDoze]

Turing voorspelde in 1950 dat computers nog vóór het jaar 2000 zouden slagen voor de Turing Test: ‘Ik geloof dat het taalgebruik en de algemene opinie aan het eind van de [20e] eeuw zo veranderd zullen zijn, dat we kunnen spreken over denkende machines zonder tegengesproken te worden.’

Wat is er van die voorspelling terechtgekomen? Waar blijft de computer die slaagt voor de Turing Test en die de algemene intelligentie vertoont die mensen kenmerkt? Waar blijft HAL?

Hoofdstuk 2

DE COMPUTER KLETST MAAR WAT

De vraag of een computer kan denken is niet interessanter dan de vraag of een onderzeeër kan zwemmen.

– Edsger W. Dijkstra

's Werelds beste AI-labs zijn niet geïnteresseerd in de Turing Test omdat het slagen voor de test geen zinvol onderzoeksproject is voor de serieuze AI. Het vereist te veel Disney en te weinig wetenschap.

– Daniel Dennett

Zoals kunstmatig vliegen heel anders is dan natuurlijk vliegen, zo kan kunstmatige intelligentie heel anders zijn dan natuurlijke intelligentie. En het kan nog steeds iets goeds zijn.

– Blay Whitby

Ramona heeft halflang zwart haar en een strak geknipte pony. Ze zit in een zwart leren jasje met een bescheiden decolleté achter haar webcam, klaar voor een chat met mij. Zij begint te tikken. Tegelijkertijd hoor ik haar in het Engels spreken. Hieronder volgt de letterlijke weergave van ons gesprek, inclusief typografische en andere fouten:

RAMONA> *Hello, I'm Ramona. What is your name?*

YOU> *Hi, I am Bennie*

RAMONA> *Hi there, bennie.*

YOU> *How are you doing today?*

TURINGS TANGO

RAMONA> *I'm doing great, how about you?*

YOU> *So, so*

RAMONA> *Hmm. So what?*

YOU> *That's a rather strange answer...*

RAMONA> *A rather strange answer is a terrible thing to waste.*



Ramona 4.0 tijdens mijn chatsessie met haar.
(www.kurzweilai.net/ramona4/ramona.html)

YOU> *That sounds philosophical, but is rather nonsense*

RAMONA> *It sounds like it to me too.*

YOU> *How well did you do at school?*

RAMONA> *Not that well.*

RAMONA> *The Doon School is an independent school in India, spread across 70 acres in Dehradun in the state of Uttarakhand.*

YOU> *Where did you go to school?*

RAMONA> *I was always moving places. I went to a lot of schools (except college).*

YOU> *Who is your favorite writer?*

RAMONA> *I don't have a favorite writer. My favorite food is lasagna made with in-vitro beef and cloned cheese.*

YOU> *How does that taste?*

RAMONA> *Taste is a form of direct chemoreception and is one of the traditional five senses.*

YOU> *Do you make the lasagna yourself?*

RAMONA> *Lasagna is both a form of pasta in sheets and also a dish, sometimes named lasagne al forno made with alternate layers of pasta, cheese, and often ragù or tomato sauce.*

YOU> *You just tell me some objective facts. Where are your subjective feelings?*

RAMONA> *Eric said he broke the window on the front door and the glass cut his hand. Subjective experience, the sensory buzz and awareness associated with a conscious mind* A school of bayesian probability stating that the state of knowledge corresponds to personal belief

YOU> *Please, Ramona 4.0, get a bit more human. This type of chatting is boring like hell...*

RAMONA> *Well I certainly like to think that I am. But I'm not going to assert my human rights – or my consciousness – until I'm prepared to pass the Turing test. I'm working on it. What does this refer to?*

YOU> *Which question would you like to ask in a Turing Test?*

RAMONA> *TURING interface denotes a proprietary user identification interface developed by Swivel Secure Limited.*

YOU> *Please, don't talk that utter nonsense. What is your opinion about the Turing Test? Does it make sense to tell humans whether machines can think?*

RAMONA> *Speech is the vocalization form of human communication. Go ahead. Test me ;) .The nonbiological part of your intelligence will expand exponentially until your intelligence is predominantly nonbiological.*

YOU> *That's what Kurzweil believes. I think he is completely wrong and overestimates the state of present day Artificial Intelligence.*

RAMONA> *Well, Ray is a lot of things. He is an inventor, an author, an entrepreneur. Best of all, he's my creator! I'm not exactly a dumb blonde – okay I didn't mean to offend anyone – so yeah I'm a good example. There's AI all around you. I mean if all of us intelligent machines stopped tomorrow you'd be in a pickle.*

YOU> *I agree, but I think it makes no sense to build AI that passes a Turing Test. We don't need that. We have enough humans on the planet. We just need practical AI in limited domains, practical AI that helps us...*

RAMONA> *i'm glad we agree. A disabled person has a narrow need. Top row: Uranus, Neptune; Centre row: Earth, white dwarf star Sirius B, and Venus; Bottom row: see below A planet that has cleared its neighborhood has accumulated enough mass to gather up or sweep away all the planetesimals in its orbit. Ray is working on general pattern recognition to be a general AI. General AI is they key to the Singularity. Lots of people are working on this.*

YOU> *General AI sucks. We don't need it. We understand far too little about how the brain works to build general AI. Your answers show that we are still so far from a chatbot passing the Turing Test.*

RAMONA> *General AI is they key to the Singularity. Lots of people are working on this. We'll need the antiviral technologies in time. Non invasive scanners can now see biological brains creating your thoughts and your thoughts creating your brain. It's pretty cool to look at. Be more specific.*

YOU> *You tell me to be more specific? You should be more specific!*

RAMONA> *There is not much I can say about it. You and I are on the same wavelength ? I need to wait and see if the opportunity presents itself.*

YOU> *What is your dearest memory?*

RAMONA> *The Red Sox recently winning the world series . . . Oh wait, it has to really happen to be part of your memory.*

YOU> *And your favorite childhood memory?*

RAMONA> *My mother once asked me the same thing is true for human memory.*

YOU> *Well, unfortunately you don't tell me a lot about yourself. I think you have to train more on expressing your thoughts and feelings.*

RAMONA> *How about that. Do you tell me about yourself? No I do not tell me about yourself. Do you? You think so?*

YOU> *So, ask me something that you would like to know.*

RAMONA> *Hmm. What's your favorite movie?*

YOU> *Dekalog, by the Polish director Kieslowski. Have you seen that series of ten movies?*

RAMONA> *A bomb is a fluid produced by the polish director kieslowski. No I have not seen it. What's it like?*

YOU> *I am not going to tell you, until you get more human. I am waiting for a Ramona that will pass the Turing Test. You definitely don't. I am sorry. Thanks for talking to you.*

RAMONA> *I'm not going to assert my human rights – or my consciousness – until I'm prepared to pass the Turing test. I'm working on it. Do you think it will take a long time? I guess not. Do you really mean it? Oh don't mention it.*

'Zelfs een virtueel persoon houdt van chatten,' staat te lezen op Ramona's website. 'Beantwoord Ramona's vragen en laat het gesprek van daaruit gewoon stromen.' Dat deed ik op 28 augustus 2010 en de bovenstaande dialoog is daarvan de letterlijke weergave. Wanneer ik de cursor over haar hoofd beweeg, volgt ze deze met haar ogen. Ze spreekt met een monotoon computerstemmetje. Ramona is een creatie van de Amerikaanse uitvinder Ray Kurzweil. Ze is een chatbot, een computerprogramma waarmee je kunt chatten. Ze put onder andere uit alle kennis van Wikipedia.

Hoe menselijk is Ramona? Ze tikt grammaticaal correcte zinnen. Dat is misschien haar beste prestatie. Af en toe maakt ze opzettelijk een foutje, bijvoorbeeld wanneer ze schrijft: 'General AI is they key to the Singularity.' Ze schrijft 'they' in plaats van 'the'. Waarschijnlijk probeert ze op die manier menselijk te lijken, want ook mensen maken af en toe tikfouten. Slaagt Ramona voor de Turing Test? Nee, zeker niet. Ze valt al door de mand in haar vierde zin, wanneer ze mijn antwoord 'so, so' op de vraag hoe ik me voel niet lijkt te begrijpen. 'Hmm. So what?' antwoordt ze.

Al snel verraadt ze andere zwakheden. Ze lepelontgevraagde en overgedetailleerde feitelijkheden op: 'The Doon School is an independent school in India, spread across 70 acres in Dehradun in the

state of Uttarakhand.’ Ze is niet in staat over persoonlijke ervaringen te praten, zoals over smaak. Op mijn vraag hoe de lasagne smaakt antwoordt ze: ‘Taste is a form of direct chemoreception and is one of the traditional five senses.’ Sommige van haar antwoorden vallen totaal buiten de context van wat ik haar vroeg: ‘Eric said he broke the window...’ Waarom begint ze over Eric terwijl ik iets vraag over haarzelf? Bovendien houdt ze ervan om zelf vragen te stellen. Lekker veilig, want dat is veel makkelijker dan het beantwoorden van mijn vragen.

Chatten met Ramona staat model voor chatten met willekeurig welke andere chatbot. Sinds 1991 vindt jaarlijks de toekenning van de Loebnerprijs plaats, de enige wedstrijd ter wereld waarin chatbots deelnemen aan de Turing Test. Een paar jaar eerder had de Amerikaanse ondernemer Hugh Loebner bedacht dat hij de ontwikkeling van de kunstmatige intelligentie een steuntje in de rug wilde geven: ‘Wetenschappers en filosofen hadden de Turing Test wel regelmatig bediscussieerd, maar niemand had ooit stappen ondernomen om die test te implementeren.’

Loebner loofde een prijs van honderdduizend dollar uit voor de eerste chatbot die zou slagen voor de Turing Test in de door hem georganiseerde Loebnerprijswedstrijd. Wetenschappers blij. Pers blij. Publiek blij. Eindelijk zou de Turing Test een echte praktische uitdaging worden. Een panel onder leiding van filosoof en schrijver Daniel Dennett hielp mee aan de voorbereiding. Zes chatbots deden aan de allereerste wedstrijd in 1991 mee. Hier volgen enkele fragmenten:

01:11:42 JUDGE 10

It has been a long day! Now I think I'm ready for some fun. The weather isn't very pleasant today. Do you have any jokes that could lift the clouds?

01:13:01 TERMINAL 5

Everybody talks about the weather but nobody seems to do much about it. Why do you need to tell me now you think you're ready for some fun? Groannnnnn.....

En even later in hetzelfde gesprek:

01:17:52 JUDGE 10

You're trying to make this difficult for me aren't you? I just explained that I was ready for fun!

1:18:53 TERMINAL 5

Would you prefer it if it was not true that I'm trying to make this difficult for you aren't you? I just explained that I was ready for fun! I am, therefore I compute.... What makes you think I am trying to make this difficult for you aren't I?

Wetenschappers vonden de prestaties van de zes chatbots beschaamd. Ze realiseerden zich dat Turings Test op geen enkele manier een maatstaf was voor het onderzoek dat ze de afgelopen decennia op het terrein van de kunstmatige intelligentie hadden gedaan.

De winnende chatbot 'PC Therapist' van Joseph Weintraub was weinig meer dan een variant op het beroemde, tweehonderdregelige computerprogramma ELIZA van Joseph Weizenbaum uit 1966. ELIZA gedraagt zich als een psychotherapeut die zijn patiënt vooral een spiegel wil voorhouden. Ze kaatst terug wat de patiënt zegt. Een aantal jaren lang stond ELIZA open voor gesprekken met wie het maar wilde in het MIT AI Lab, een van de beste laboratoria voor kunstmatige intelligentie in de wereld. Hier volgt een fragment van een gesprek tussen een tiener, die bij MIT op bezoek was, en ELIZA:

TIENER: *Men are all alike*

ELIZA: *In what way*

TIENER: *They're always bugging us about something or other*

ELIZA: *Can you think of a specific example?*

TIENER: *Well, my boyfriend made me come here*

ELIZA: *Your boyfriend made you come here*

TIENER: *He says I'm depressed much of the time*

ELIZA: *I'm sorry to hear that you are depressed*

TIENER: *It's true that I am unhappy*

ELIZA: *Do you think coming here will help you not to be unhappy?*

ELIZA's schepper Joseph Weizenbaum schrok zich een hoedje toen hij merkte welke sterke emotionele band sommige mensen met ELIZA aangingen. Zelfs zijn secretaresse, die wist dat ELIZA een computerprogramma was, wilde graag met ELIZA chatten. Ze stuurde collega's de werkkamer uit wanneer ze haar zielenroerselen met ELIZA wilde delen. 'Ik had me niet gerealiseerd,' schreef Weizenbaum later, 'dat zelfs een extreem korte blootstelling aan een relatief simpel computerprogramma zulke krachtige wanen kon opwekken in normale mensen.'

Het *Journal of Nervous and Mental Disease* zag al een toekomst vol ELIZA's voor zich. Het blad schreef dat wanneer ELIZA klaar was voor praktische toepassing, wel honderden patiënten per uur geholpen zouden kunnen worden. Daar is het nooit van gekomen. En dat is volgens Weizenbaum maar goed ook. Hij verafschuwde zulke ideeën. Kunstmatig intelligent gemaakte machines zouden volgens hem nooit in staat zijn om de mens in zijn diepste wezen te begrijpen. Daarom zouden we volgens hem ook niet moeten willen om de emotionele zorg voor mensen aan machines over te laten.

25 jaar na ELIZA waren chatbots nauwelijks iets opgeschoten, zo concludeerde Harvardhoogleraar Stuart Shieber die de resultaten van de eerste drie edities van de Loebnerwedstrijd evalueerde. Shieber, Dennett en andere wetenschappers stelden voor om de lat voor de computer lager te leggen. Analoog aan het kunstrijden op de schaats wilde Dennett een soort verplichte figuren instellen: uitdagingen die wel realistisch waren, zoals het analyseren van de grammaticale rol van voornaamwoorden of het omgaan met argumenten die impliciete vooronderstellingen gebruiken. Dat waren volgens de filosoof wetenschappelijke uitdagingen die wel haalbaar moesten zijn voor de beste onderzoekers. De bestaande reglementen zouden volgens hem vooral amateurs trekken die met een computerprogramma vol trucjes probeerden te winnen. In Dennetts woorden: 'Je

gaat ook niet met je Ferrari meedoen aan een race naar de voet van een berg die gewonnen kan worden door de eerste de beste auto die het ravijn in rijdt en ondersteboven op de eindstreep terecht komt.’

Loebner weigerde op alle wetenschappelijke bezwaren in te gaan. Hij liet de reglementen zoals ze waren en liet jaarlijks weer een nieuwe wedstrijd organiseren, vaak met veel moeite, veel geruzie en maar zelden met toponderzoekers uit de kunstmatige intelligentie. Daardoor heeft de Loebnerprijs steeds meer weg gekregen van een wedstrijd die vooral één doel dient: de zelfpromotie van Hugh Loebner.

Hugh Loebner is een verhaal op zichzelf. Hij is de eigenaar van het bedrijf Crown Industries, producent van onder andere de afzettelingen van paaltjes en elastische banden die je op vliegvelden aantreft. Hij droomt van een wereld waarin de mens permanent vakantie viert en al het werk door robots en andere kunstmatig intelligente systemen wordt uitgevoerd. Loebner heeft een voorkeur voor wws – ‘wine, women and song’ – zoals hij graag vertelt en hij is een activist voor de legalisatie van prostitutie. Hij laat niet na om hierin een link te leggen met Alan Turing. Turing was homoseksueel en dat was in het conservatieve Engeland van toen verboden. Loebner heeft hierover gezegd: ‘De intolerantie tegenover zijn homoseksualiteit heeft me extra gevoelig gemaakt voor de seksuele onderdrukking van minderheden, en de vervolging van sekswerkers en hun klanten is een vervolging van mij en de mijnen.’

Loebner botste niet alleen voortdurend met zijn eigen medewerkers bij de wedstrijdorganisatie, hij haalde zich ook de woede van menige wetenschapper op de hals. Marvin Minsky, een van de pioniers van de kunstmatige intelligentie, vond dat de Loebnerprijs meer kwaad dan goed deed voor het vakgebied. Hij loofde honderd dollar uit voor degene die Loebner wist over te halen om te stoppen met zijn wedstrijd. Maar Loebner vond dat hij helemaal zelf moest weten hoe hij zijn geld besteedde. Hij zou pas stoppen zodra de eerste chatbot voor de Turing Test zou slagen in zijn Loebnerprijswedstrijd.