

Ray Kurzweil – The Age of Spiritual Machines

Twee universele wetten

Kurzweil begint zijn boek met een filosofische ondersteuning voor zijn toekomstvoorspellingen. Hij introduceert zijn *wet van de versnellende opbrengsten* (*law of accelerating returns*). Hoe verder de orde toeneemt, hoe sneller de tijd gaat. De belangrijkste vorm van orde waar Kurzweil het in dit boek over heeft is de steeds groter wordende orde die de mens in de wereld schept door de technologie steeds verder te ontwikkelen. Met het sneller

gaan van tijd bedoelt hij dat belangrijke gebeurtenissen elkaar in een steeds sneller tempo opvolgen (gemeten in absolute tijd). De exponentiële toename van de orde in de wereld zoals die nu aan de gang is, zorgt volgens deze wet voor een versnelling van de tijd.

De *wet van de toenemende chaos* (*law of increasing chaos*) houdt precies het omgekeerde in: als de chaos exponentieel toeneemt, zal de snelheid van de tijd exponentieel afnemen. Neem bijvoorbeeld het ontstaan van het heelal, waar al na 10^{43} seconde zwaartekracht ontstond, en andere belangrijke gebeurtenissen volgden na telkens exponentieel langere perioden. Zo wordt er bijvoorbeeld na 1 seconde, na 1 minuut en vervolgens pas na 300.000 jaar een mijlpaal in de geschiedenis van het heelal bereikt. Dit wordt veroorzaakt door een toenemende chaos: het heelal begon als een tamelijk egaal, over een kleine ruimte verdeelde, hete brei en eindigt in al haar complexiteit als ‘betoverd’ van miljarden sterrenstelsels.

Hoewel deze wetten inderdaad voor veel dingen opgaan, heb ik toch twijfels bij de universaliteit (en dus de wetenschappelijke waarde) ervan. Neem bijvoorbeeld een brok graniet dat deel uitmaakt van een berg. Deze kan uit heel veel dezelfde moleculen en structuren bestaan en dus heel geordend zijn. Toch is de kans groot dat er eeuwenlang vrijwel niets aan deze steen verandert. Stel dat deze steen op een gegeven moment toch toevallig verandert: hij komt bij een aardbeving aan de (chaotische) oppervlakte van de berg, breekt en wordt vervuild. Hierdoor wordt hij chaotischer. Er is geen enkele reden om aan te nemen dat deze steen vervolgens minder zou veranderen dan toen hij nog veilig in zijn berg zat.

Evolutie en technologie

Kurzweil ziet een gelijkheid tussen de evolutie en de ontwikkeling van de technologie. De evolutie heeft als resultaat van zijn jarenlange probeersels de mens als resultaat gehad. Je zou kunnen zeggen dat de evolutie een bepaald soort intelligentie bezit, zij het van een vorm die voor mensen makkelijk te overzien is, namelijk door

Een eerste aanblik van dit boek doet niet veel serieus vermoeden. De zilveren kافت straalt je in alle kleuren van de regenboog toe. Als je het boek openslaat blijkt de auteur net in een gesprek verwickeld te zijn met een vriendin van hem die in het jaar 2099 leeft. Toch is dit boek te vinden in de universiteitsbibliotheek en is het geschreven door een vooraanstaande Amerikaanse AI'er. Hij heeft in de afgelopen decennia een aantal belangrijke prijzen in de wacht gesleept, waaronder één voor de voorganger van dit boek, *The Age of Intelligent Machines*. Halverwege het schrijven van zijn boek is hij overgegaan van typen naar het mondeling dicteren van zijn teksten aan software die in zijn eigen bedrijf ontwikkeld is.

botweg willekeurig proberen (d.w.z. combineren en muteren van DNA) en maar zien of er een keer iets beters uitkomt. Zoals de ‘domme’ evolutie de ‘slimme’ mens geschapen heeft, zal de ‘domme’ mens op zijn beurt de ‘slimme’ techniek voortbrengen. Deze techniek zal uiteindelijk de intelligentie van mensen ver overstijgen. En omdat de mens veel geordender werkt, zal, volgens de wet van de versnellende opbrengsten, deze ontwikkeling veel sneller gaan. Waar de evolutie nog miljarden jaren nodig heeft, heeft de mens nog maar zo'n 100.000 jaar nodig. (Dat is de duur van het bestaan van de mensheid tot nu toe.) Volgens Kurzweil zullen machines namelijk al voor het einde van deze

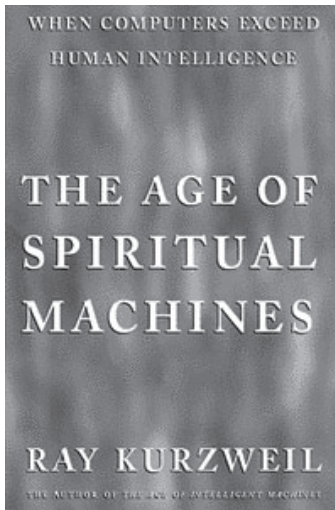
eeuw alle menselijke aspecten van intelligentie bezitten.

Angstaanjagend?

Als we veel sciencefictionfilms mogen geloven gaan intelligente machines met een kwade wil de mensheid met grof geweld overnemen (wat in Hollywood natuurlijk toch net niet lukt). Zo gaat het volgens Kurzweil niet.

Op dit moment kunnen enkele menselijke taken veel beter door machines uitgevoerd worden dan door mensen, zoals rekenen en letterlijk dingen onthouden. Die set taken zal zich gestaag uitbreiden de komende jaren. Exponentieel snel zelfs, volgens de eerder beschreven wetten. We zullen de komende decennia geconfronteerd worden met veel implantaattechnologie waarmee we onze zintuigen, ons redeneervermogen en ons geheugen kunnen opkrikken. De mensen die dit soort technologie uit wantrouwen of angst niet in hun lichaam willen toelaten, zullen achterblijven bij hun steeds competentere omgeving. En mensen die bang zijn om achter te blijven, zullen er toch aan moeten. Het begrip ‘mens’ en ‘machine’ zullen langzamerhand in elkaar gaan overlopen: mensen krijgen steeds meer machinale componenten en machines worden steeds menselijker.

Uiteindelijk zullen MRI-hersenscans zo nauwkeurig zijn dat neuronen en hun afzonderlijke componenten één voor één gedetecteerd kunnen worden. Aangezien onze intelligentie niet veel meer is dan deze miljarden neuronen die op een bepaalde manier met elkaar verbonden zijn, zullen we dan in staat zijn onze volledige intelligentie te scannen. De processorkracht is tegen die tijd groot genoeg om met die gegevens aan de slag te kunnen, en daarmee de hersenprocessen kunstmatig door te zetten. Resultaat is dat een mens, met zijn complete persoonlijkheid, ervaringen en intelligentie, in een computer voortleeft. We zullen in staat zijn een effectieve kopie van de gescande persoon te produceren. De relatie tussen het origineel en zijn kopie (of kopieën) is op zijn zachtst



Uitgave: Viking Penguin, 1999
ISBN: 0-670-88217-8

heeft Kurzweil hier een oplossing voor, namelijk de enorm verbeterde *virtual reality* die tegen die tijd beschikbaar zal zijn. Door implantaten in het ruggenmerg van niet-gescande mensen zal men zich al in 2029 in volledig andere, realistische ervaringen kunnen wanen. De ervaring die wij hebben van ons lichaam is slechts gebaseerd op zenuwpulsen die in onze hersenen grotendeels via het ruggenmerg binnenkomen, dus die ervaringen zijn volledig te simuleren mits computers voldoende rekenkracht hebben om een realistische virtuele omgeving om te zetten in zenuwpulsen. The Matrix, de film waarvan het eerste deel in hetzelfde jaar uitkwam als dit boek (1999), kan werkelijkheid worden. En als we implantaten kunnen maken die deze *virtual reality* kunnen produceren, is het natuurlijk een eitje om die *virtual reality* aan een ingescand mens te koppelen. Dat deze personen alleen nog in *virtual reality* met een gewoon lichaam kunnen leven en bewegen, is geen probleem aangezien ook mensen met een ouderwets lichaam het grootste deel van hun tijd in *virtual reality* zullen doorbrengen. Tegen deze tijd is het wel tegen 2099, en zal het onderscheid en de relatie tussen allerlei verschillende werkelijkheden moeilijk te begrijpen zijn. Maar hier hoeven we niet angstig over te zijn: onze toegenomen intelligentie zal het ongetwijfeld makkelijk kunnen bijhouden.

Mensen en machines hoeven dus niet te concurreren zoals in sf-films: mensen worden gewoon machines, en andersom. Ondanks de vele morele bezwaren die bij deze nieuwe techniek komt kijken, biedt zij grote voordelen: de mens kan eindeloos voortleven zolang hij niet door een externe kracht 'gewist' wordt. Computervirussen zullen de belangrijkste bedreiging zijn van de gezondheid in het jaar 2099.

Wat staat ons te wachten?

Of de AI daadwerkelijk zo'n vlucht zal nemen is natuurlijk nog maar zeer de vraag. Kurzweil voorspelt de meest bizarre dingen, terwijl niemand met zekerheid kan zeggen of de technologie zich ooit zo ver zal ontwikkelen. Wat ik tijdens het lezen van dit boek gemist heb is de onderbouwing van zijn theorie dat er ruim voldoende materie in onze omgeving is om computers de komende eeuw(en) van intelligentie te voorzien. Hij betoogt tegen het einde van zijn boek dat het grootste deel van de materie op aarde niet

gezegd vreemd, maar filosofisch erg interessant. Kurzweil geeft ook een scenario waarin een mens destructief gescand moet worden, waarin het lichaam van een persoon kapotgemaakt moet worden teneinde hem in te scannen. Heb je dan die persoon vermoord of alleen zijn lichaam?

Over lichamen gesproken: ingescande mensen zitten natuurlijk wel zonder. Gelukkig

gebruikt wordt voor intelligente processen, wat natuurlijk waar is, en dat er dus nog heel veel mogelijk is met de materie die er is, grenzeloos veel zelfs. De onderbouwing hiervan laat hij liggen tot een volgend boek, wat hij tot nu toe helaas nog niet geschreven heeft. Kurzweil plaatst in de epiloog ook nog een aantal kanttekeningen bij zijn betoog, bijvoorbeeld het realistische scenario dat de mensheid zichzelf zal vernietigen.

Dat Kurzweil verder goed aan zijn argumentatie gewerkt heeft moet ik direct toegeven. Aan de ene kant baseert hij zijn theorieën op een uitgebreide kennis van het verleden en zijn eigen ervaring met geschiedenis schrijven als AI'er. De technologie is de afgelopen 100 jaar in een steeds sneller en verbazingwekkender tempo vooruitgegaan. Er worden inderdaad in steeds sneller tempo nieuwe dingen bedacht en gemaakt op dit moment, en dit zal voorlopig nog wel even doorgaan. Wie had bijvoorbeeld 5 jaar geleden gedacht dat je nu raar aangekeken wordt als je niet de hele dag met je mobieltje rondloopt? Zelfs als de ontwikkeling van de techniek niet exponentieel versnelt, staat ons nog heel wat te wachten de komende eeuw. En als AI'ers gaan we die toekomst nog zelf scheppen ook.

Tot slot

Dat Kurzweil zijn eigen voorspellingen ook niet helemaal (wetenschappelijk) serieus neemt blijkt wel uit zijn manier van schrijven. Aan het einde van elk hoofdstuk wordt hij ondervraagd door een persoon, die 'de lezer' blijkt te zijn. Deze ondervraagt Ray over een aantal dingen die hij in dat hoofdstuk heeft uitgelegd, wat af en toe best geestige dialogen oplevert en bovendien met een frisse andere blik tegen de zaken aankijkt. In de loop van het boek blijkt deze persoon Molly te heten, en als Kurzweil met zijn toekomstvoorspellingen begint, gaat Molly mee met de tijd terwijl Ray achterblijft in 1999. De rollen worden nu omgedraaid: Kurzweil vraagt aan Molly hoe het er aan toegaat in de toekomst. Molly krijgt kinderen, trouwt, krijgt machinale implantaten, scheidt van haar menselijke man, trouwt vervolgens met een machine van wie ze meer houdt en leeft in 2099 nog steeds machinaal voort op 120-jarige leeftijd. Tegen die tijd is de complexiteit van de verschillende realiteiten waarin mensen (tegelijktijd) dingen aan het doen zijn haast niet meer te volgen voor een ouderwets mens met biologische hersenen. Kurzweil geeft verder een mooi overzicht van de geschiedenis van de technologie en de AI die daaruit voort is gekomen, zij het dat deze informatie natuurlijk exponentieel snel ouder wordt: 6 jaar is tegenwoordig een eeuwigheid. Hij beschrijft verder hoe je een intelligente machine kan maken door middel van neurale, recursieve of evolutionaire algoritmen, of natuurlijk een combinatie van die drie, maar meer dan een basale uitleg voor leken geeft hij niet. Het lijkt me voor de net beginnende eerstejaars AI'er een goed boek om een beetje bekend te raken met de basisprincipes van de AI en de filosofische en ethische dilemma's waar je als AI'er mee te maken kan krijgen. En natuurlijk om erachter te komen hoe enorm spannend AI kan zijn. ☺