

1 De Turing test

De Turing test stelt dat een machine geacht kan worden intelligent te zijn en te denken wanneer een jury een menselijke gesprekspartner niet kan onderscheiden van de machine in een dubbel blinde test (Turing 1950).

1.1 Als herformulering van de vraag “kunnen machines denken?”

De Turing test wordt gepresenteerd als operationalisering voor de vraag “kunnen machines denken?”. Ik ben van mening dat de test niet goed aansluit bij deze vraag, omdat de test puur behaviouristisch is. Er wordt enkel gekeken naar het gedrag van de computer, niet naar de werking ervan.

Door het definiëren van “denken” te ontwijken presenteert Turing zich als een behaviourist *pur sang*. Zelf zie ik twee mogelijkheden om “denken” te definiëren:

1. Dat wat mensen doen (in tegenstelling tot dieren en levenloze objecten)
2. Dat wat de menselijke geest doet (en mogelijk die van dieren), inclusief de manier waarop dit gebeurt.

Turing neigt duidelijk naar de eerste definitie, en dat is tevens wat de test onderzoekt. De test past echter niet bij de tweede definitie, tenzij men de vraag omdraait, en zich afvraagt of mensen misschien zelf machines zijn? Turing zag intelligentie als geidealiseerde rationaliteit, mogelijk door zijn persoonlijke aversie voor hypocrisie en irrationaliteit.¹

Deze interpretatie is natuurlijk niet noodzakelijkerwijs verbonden met de Turing test. Men kan in de eerste definitie geloven, en aannemen dat een computer het gedrag van een mens kan simuleren, op een voor de computer passende wijze. Echter, om dan de Turing test te gebruiken staat gelijk aan het verwerpen van de wet van het behoud van energie omdat een illusionist zijn assistent schijnt te laten verdwijnen (ook al bestaat deze illusie slechts voor het publiek).

Nee, wanneer er buitengewone beweringen worden gedaan [iemand verdwijnt, een machine is intelligent], dan zijn buitengewone verklaringen benodigd.² Ongebrijdelde skepsis is geboden! De Turing test vormt echter een obstructie in deze. Het ongeloof dient tijdelijk te worden afgezworen, men dient met beperkte informatie te beslissen over zulk een fundamentele vraag als die van machinale intelligentie.

Een voorbeeld van een buitengewone claim die bij nader inzien verworpen moest worden is “der kluge Hans”³, een paard dat zou kunnen rekenen, aanvankelijk beschouwd als een bewijs voor de intelligentie van paarden. Bij nader inzien bleek echter dat zijn baas, overigens zonder opzet, het paard met lichaamstaal signaleerde dat hij het goede antwoord gegeven had. Dit is naar mijn mening een aanwijzing dat met alles rekening moet worden gehouden. Dit kan men zien als een argument tegen de Turing test, omdat deze intelligentie reduceert tot het maken van een indruk zonder deze nader te bestuderen.

Verder is het vrij arbitrair om alleen naar taaluitingen te kijken, en daarmee andere vormen van intelligentie uit te sluiten. Het lijkt erop alsof de hersenen bij mensen in twee helften zijn verdeeld, waarbij de een zich vooral met taal bezighoudt, en de ander met niet talige aspecten. Er zijn nog veel niet

¹Hodges (1992) schrijft:

“[...] In fact his idea of society was that of an aggregate of individuals, much closer to the views of democratic individualism held by J. S. Mill than that of socialists. And to keep his individual self intact, self-contained, self-sufficient, uncontaminated by compromise or hypocrisy* was his ideal.”

²Parafrase van:

“And when such claims are extraordinary, that is, revolutionary in their implications for established scientific generalizations already accumulated and verified, we must demand extraordinary proof.” – Editorial in *The Zetetic* (Vol. 1, No.1, Fall/Winter 1976, p 4)

³Zie http://en.wikipedia.org/wiki/Clever_Hans

talige handelingen waar computers moeite mee hebben, zoals het herkennen van objecten, het plannen van bewegingen en het werken met patronen. Er is geen reden om aan te nemen dat deze taken minder afhankelijk zijn van intelligentie dan het gebruik van taal.

Gegeven een grote hoeveelheid data, bijvoorbeeld transcripties van gesprekken over onderwerpen waar de juryleden bekend mee zijn, kan het mogelijk zijn om de jury lang genoeg te overbluffen. Vergelijk het aanpassen van Deep Blue tussen partijen, toen grootmeester Gary Kasparov werd verslagen⁴ (Hamilton & Garber 1997). Er werden strategieën toegevoegd aan Deep Blue om (door mensen!) geobserveerde zwakheden van Kasparov te exploiteren. Op eenzelfde wijze zou een “intelligente” machine incrementeel verbeterd kunnen worden totdat deze over genoeg informatie beschikt om de jury om de tuin te leiden – naast de gebruikelijke methodes zoals ontwijkende antwoorden en wedervragen wanneer er geen repliek valt te matchen bij een uiting (Krol 1999).

Als laatste tegenwerping is er nog het paradoxale karakter van de test. Stel dat een computer glansrijk slaagt voor de Turing test, dan heeft de jury kennelijk besloten dat de computer menselijker is [overkomt] dan een mens — maar kan iets menselijker zijn dan een mens? Echter dan echt doch nagemaakt?⁵

Toch kan de test nuttig zijn, het vergt namelijk een bijzonder goed programma om een jury te misleiden. Maar de test moet naar mijn mening op zichzelf staan, en niet worden geïdentificeerd met de vraag of machines kunnen denken. De vraag is eerder, “kan deze machine mensen voor de gek houden?”. De uitkomst zegt naar mijn mening even veel over de mensen (die overtuigd worden) als over de computer, aangezien de jury bestaat uit mensen, die een subjectief oordeel zullen vellen.

1.2 Tegenwerpingen op de test

Als tegenwerpingen tegen de test noemt Turing dat het voor mensen moeilijk zal zijn zich voor te doen als computers, doordat bijvoorbeeld rekenvragen met grote getallen voor een computer geen problemen horen te vormen. Verder noemt hij dat het voor de computer misschien niet de beste strategie zal zijn om zich als man voor te doen, maar hier acht hij het niet nodig daar verder op in te gaan.

Turing noemt deze tegenwerpingen belangrijk, maar noemt geen oplossingen, vermoedelijk omdat het slechts om details van de spelregels gaat.

1.3 Argument voor

Turing geeft geen specifiek argument voor zijn vertrouwen dat computers voor de test zullen slagen. Het is eerder een gebrek aan (voor hem) steekhoudende tegenargumenten dat hem optimisme verschaft. Hij neemt aan dat het vooral een kwestie van veel programmeren zal zijn. Sinds de publicatie van zijn artikel zijn er echter verscheidene theoretische en praktische problemen ontdekt bij het nabootsen van menselijke intelligentie (zoals bijvoorbeeld het frame probleem, symbol grounding probleem, het parallele karakter van het brein vs. computers met slechts enkele processors, common sense knowledge etc.)

1.4 Tegenargument

Het “argument van bewustzijn” vind ik de sterkste tegenwerping tegen het geloof dat machines zullen kunnen denken. Het klopt dat mensen onderling de “beleefde gewoonte” hebben om aan te nemen dat andere mensen een bewustzijn hebben, maar dat betekent nog niet dat die gewoonte nergens op gebaseerd is. Mensen zijn geëvolueerd om levende dieren van objecten te onderscheiden. Verder worden mensen al gauw achterdochtig als iemand wartaal uitslaat, want dan moet worden aangenomen dat er iets mankeert aan het bewustzijn van de persoon in kwestie.

Turing labelt deze tegenwerping als “solipsisme”; dit is echter onjuist, het betreft het “other minds problem” (hoewel deze twee problemen verder filosofisch gezien even onoplosbaar zijn) (Harnad 2004).

1.5 Chatbots

De huidige chatbots slagen nog niet voor de Turing test (Krol 1999). Dit komt omdat het nog niet altijd lukt om een relevant antwoord te geven op vragen van mensen. Verder is het vrij moeilijk om een

⁴Dat dit mocht was overigens vooraf overeengekomen, wellicht had Kasparov hier beter niet mee kunnen instemmen.

⁵Dieren reageren soms op supranormale prikkels — zo kan een roodborstje een object met een felrode stip voor een ouder aannemen; het zou echter een sterk staaltje behaviourisme zijn om te beweren dat menselijke intelligentie instinctief herkend wordt!

gespreksonderwerp te infereren uit een reeks uitspraken (de *discourse*); de computer geeft zich te kennen door incoherent over te komen. Ten laatste, en tevens het belangrijkste, er is een ongekende hoeveelheid kennis over de wereld nodig om op simpele, *common sense* vragen te kunnen antwoorden. Aangezien de chatbots nog geen fysieke belichaming hebben kunnen ze deze kennis ook niet empirisch vergaren, en is hier veel handwerk voor nodig.

1.6 Bronnen

A.M. Turing, 'Computing Machinery and Intelligence', *Mind* LIX, no. 2236 (Oct. 1950): 433-60
<http://cogprints.org/499/00/turing.html>

Stevan Harnad (2004) The Annotation Game: On Turing (1950) on Computing, Machinery, and Intelligence, in Epstein, Robert and Peters, Grace, Eds. *The Turing Test Sourcebook: Philosophical and Methodological Issues in the Quest for the Thinking Computer*. Kluwer.
<http://cogprints.org/3322/>

Andrew Hodges (1992), *Alan Turing: The Enigma*. Vintage, Random House.
Voor het geciteerde zie: <http://www.turing.org.uk/book/extracts/ext2-70.html>

Scott Hamilton, Lee Garber, "Deep Blue's Hardware-Software Synergy," *Computer*, vol. 30, no. 10, pp. 29-35, October, 1997.

Marina Krol, "Have We Witnessed a Real-Life Turing Test?," *Computer*, vol. 32, no. 3, pp. 27-30, March, 1999.