

Können Maschinen denken?

Animals, Zombanimals und der Totale Turingtest

Markus Mottl*

4. Dezember 2002

Zusammenfassung

In ihrer Publikation „Animals, Zombanimals, and the Total Turing Test“¹ versuchen Selmer Bringsjord et al. zu zeigen, dass der „totale Turingtest“, der die Überprüfung intensionaler Eigenschaften der Testteilnehmer erlaubt und der auch in bezug auf bestimmte Vergleichssubjekte, nämlich (teils imaginäre) Tiere, spezialisiert wird, ein besserer Gradmesser für Künstliche Intelligenz sei als der traditionelle Turingtest. In dieser Besprechung genannter Arbeit versuchen wir zu zeigen, dass die im Artikel dargestellten Ideen tatsächlich keine wissenschaftlich überprüfbaren Aussagen über das Vorhandensein oder die Absenz der untersuchten Eigenschaften, sofern sie existieren oder gar überhaupt sinnvoll sind, machen und darüber hinaus versteckte Grundannahmen voraussetzen, die mit naturwissenschaftlichen, allgemein akzeptierten Erkenntnissen nicht im Einklang stehen.

1 Einführung

Es ist eigentlich nicht ganz klar, was Bringsjord et al. in ihrer Arbeit zeigen wollen. Zwar behaupten sie in der Übersicht, einen besseren Gradmesser für Künstliche Intelligenz zu entwickeln, der auf dem *totalen Turingtest* aufbaut, doch schon am Ende der Einleitung wird dieses Ziel geändert: da ist auf einmal zu lesen, dass die Arbeit „ein Versuch wäre, die Idee zu präsentieren und teilweise zu verteidigen, dass KI, so wie sie zur Zeit betrieben wird, an Werkzeuge und Techniken gebunden ist, die gerademal ausreichen, *Zombanimals* zu gewährleisten, aber keine *Personen*“. Dies, nämlich ein Angriff gegen die *starke KI*, ist offensichtlich ein gänzlich höherer Anspruch! Darüberhinaus scheint hier implizit das Konzept *Zombanimal* mit *Artificial Animal*, wie in der Übersicht verwendet, gleichgesetzt zu werden, was bis zu diesem Zeitpunkt noch nicht bewiesen ist und daher eher verwirrend auf den Leser wirkt. Wie wir gegen Ende sehen werden, dehnen die Autoren den Begriff dann sogar allgemein auf Tiere (exkl. den Menschen) aus. Vielleicht sollte der Artikel überhaupt heißen: „Warum nur Menschen ein Bewusstsein haben“...

*Matrikelnummer: 9303956

¹[BCN00]

2 Von Zombies zu Zombianimals

Die Abgrenzung zum Turingtest, nämlich, dass dieser *obskure Konzepte* wie *Denken, Bewusstsein, freien Willen* und *Kreativität* „verbannt“ und für diesen die einzig interessante Frage wäre, wie ein Ingenieur konkrete Artefakte erschaffen kann, erscheint schon fast suggestiv: „verbannt“ werden nämlich bei Turing diese Konzepte nicht deswegen, weil sie, wie vielleicht die Autoren des Artikels suggerieren wollen, nicht nachgebaut werden könnten und man sich daher diesem Problem mittels Vogel-Strauss-Taktik entziehen würde, sondern schlicht und ergreifend, weil sie, linguistisch betrachtet, problematisch sind. D.h. sie sind entweder nicht wohldefiniert (mehrere, teils widersprüchliche Definitionen existieren) oder zu vage (man kann die Messlatte immer wieder verschieben, z.B. bei Kreativität).

In den darauf folgenden Absätzen wird der Begriff des philosophischen Zombies definiert, nämlich ein Geschöpf, das sich zwar nach außen hin exakt genauso verhält wie ein vorgegebenes „bewusstes“, aber im Gegensatz dazu nicht über den geringsten Funken *phänomenalen Bewusstseins* verfügt, wobei letzterer Begriff durch Eigenschaften wie *Selbstwahrnehmung, innere Gefühle, „eine Ansicht haben“*, *„was es bedeutet, etwas zu sein“* beschrieben wird.

Logisch möglich, also denkbar, ist es natürlich durchaus, dass solche Geschöpfe existieren. Allerdings ist auch denkbar, dass die Naturgesetze ganz anders sein könnten, als wir sie wahrnehmen oder dass es mehr Gesetzmäßigkeiten gibt, als wir kennen oder kennen können. Daher erscheint der Teil über die *physikalische Möglichkeit* von Zombies auch etwas fragwürdig: dort wird ein Gedankenexperiment durchgeführt, in dem alle bekannten Naturgesetze eingehalten werden und in dem das „natürliche“ Gehirn einer Versuchsperson stückweise durch ein „künstliches“ ersetzt wird, wobei das *phänomenale Bewusstsein* aber in selbem Maße verschwindet.

Dieses Argument steht aber auf wackeligen Beinen: einerseits, weil es eben ein noch unbekanntes oder gar unergründbares Naturgesetz sein könnte, dass Materie *notwendigerweise* Bewusstsein hervorbringt, wobei die Form der Materie auch die Form des Bewusstseins bedingen könnte. Dies ist genausowenig beweisbar wie das Gegenteil, das die Autoren offensichtlich annehmen.

Andererseits ergeben sich tatsächlich Widersprüche, wenn man ihr Gedankenexperiment in die reale Welt überträgt: wenn das „bewusste, natürliche“ Gehirn kleinweise durch ein „unbewusstes, künstliches“ ersetzt wird und der Patient die Änderung in seinem Bewusstsein „erlebt“, dann bedeutet dies zunächst, dass diese Wahrnehmung nicht vom per definitionem „unbewussten, künstlichen“ Gehirnteil gemacht werden kann. Wenn nun aber der „bewusste, natürliche“ Teil davon erfährt, dann wohl nur, wenn die ausgetauschten Gehirnteile nicht funktional äquivalent zum originalen Teil sind — im Widerspruch zur Experimentannahme! Funktional äquivalent bedeutet nun einmal, dass man äußerlich, und das schließt hier den verbliebenen, natürlichen Gehirnteil mit ein, keinen Unterschied feststellen kann. Blicke als letzte Rettung für dieses Argument, dass der im natürlichen Teil verbliebene Geist in dem Moment den Körper nicht mehr kontrollieren kann, in dem auch nur der kleinste Gehirnteil ausgetauscht wird, was im Text des zitierten Philosophen Searle ebenfalls zum Ausdruck kommt, wenn er sagt, dass „ihr Bewusstsein langsam im Nichts verschwindet, während das extern beobachtbare Verhalten gleich bleibt.“. Eine mehr als unsinnige Annahme, wenn man davon ausgeht, dass im Experiment natürliche Ge-

hirnteile bis zum Schluss an motorischen Nerven verbleiben können! Wenn also weder der „natürliche, bewusste“ Teil einen (subjektiven) Unterschied merkt, noch die Experimentatoren einen Unterschied im Verhalten feststellen können, wer könnte dann noch davon reden, dass überhaupt ein Unterschied existiert? Kurz gesagt, die physikalische Möglichkeit von Zombies durch inkrementelle Gehirnersetzung kann bezweifelt werden.

3 Einfache Zombianimals

Im weiteren Verlauf der Arbeit beschreiben Bringsjord et al., wie man hypothetische, biologische Kreaturen mit Mitteln der Robotik nachbaut. Hier ist schon wieder einmal Kritik anzubringen, denn was unterscheidet eine *hypothetische, biologische* Kreatur von einer rein hypothetischen? Immerhin wird die Kreatur ausschließlich durch ihr Verhalten beschrieben, also nicht im mindesten auf biologischer Basis! Damit führt sich auch die Erwähnung des totalen Turingtests ad absurdum, wenn ohnehin nur Verhalten nachmodelliert wird, um ein Zombianimal zu schaffen: das reicht auch für den normalen Turingtest. Die Aussage, dass die in diesem Teil erwähnte, künstlich geschaffene Kreatur „offensichtlich kein subjektives Bewusstsein hat“ und daher ein „Zombianimal“ sei, ist absolut nicht objektiv überprüfbar, da eben subjektiv und deswegen auch nicht „offensichtlich“. Schließlich können wir nicht einmal mit Sicherheit sagen, ob andere Menschen neben uns ein Bewusstsein haben! Die darauf aufbauende Behauptung, dass komplexere Varianten der künstlichen Kreatur auch kein phänomenales Bewusstsein haben könnten, da der kleine Unterschied zum Vorgänger dieses nicht plötzlich erzeugen könne, ist ebenso unfundiert. Dass Bewusstsein inkrementell wachsen kann, sollte eigentlich jedem Menschen klar sein, der die Entwicklung von der Eizelle zum Erwachsenen durchlebt hat.

4 Von Zombianimals zur Essenz der Künstlichen Intelligenz

Nach einer Beschreibung gängiger KI-Techniken zur Modellierung von Intelligenz kommen die Autoren zum Schluss, dass auch in Robotern verwirklichte, wissensbasierte Agenten Zombies wären, also kein „genuines, inneres Leben“ hätten, dass es mit keiner Qualität des Fühlens verbunden wäre, so eine Kreatur zu sein. Dies mag einleuchtend erscheinen, wenn man annimmt, dass wissensbasierte Agenten eine fest vorgegebene Regelmenge hätten, die ihr Verhalten steuert und dass sie nur reflexartig auf ihre Umgebung reagieren. Aber wenn man davon ausgeht, dass der Agent einem ständigen Informationsstrom aus der Umgebung unterliegt und selbst bei Fehlen desselben der Agent durch Berechnungen versucht, neue Regeln aus gespeicherten Fakten zu erstellen (durch induktive Lernverfahren), dann lässt sich dieses Argument nicht mehr so ohne weiteres halten.

5 Wenn die Induktion versagt: Personheit

Wieder schlagen die Autoren sogleich mit einer unbegründbaren Behauptung zu: „Wir sind Personen, so seid ihr.“. Nun, obwohl wir das wohl von uns sagen würden, besteht kein objektiver Grund für uns, selbiges von Bringsjord et al. anzunehmen. Natürlich würden wir ihnen dies auch zugestehen, allerdings nicht aufgrund der von ihnen angedeuteten *mathematischen Induktion*, wo man sich schrittweise von einer primitiven, künstlichen Lebensform durch Anwendung von technischen Verfeinerungen an komplexere herantastet, sondern aufgrund von *empirischer Induktion*: wir verallgemeinern Beobachtungen, projizieren Selbsterfahrung in „ähnliche“ Lebensformen. Doch diese Art von Induktion ist logisch nicht notwendigerweise konsistent, daher auch keine Garantie für Korrektheit!

Auch zeigt sich in diesem Teil der Arbeit ein Widerspruch zu einem anderen: hier wird (richtigerweise) behauptet, dass wir Menschen (*homo sapiens sapiens*) in einer realen Weise biologische Kreaturen, ja Tiere seien. Dies widerspricht allerdings später getroffenen Aussagen über die biologische Evolution, auf welche wir weiter unten noch genauer eingehen werden: Bringsjord meint dort, dass Evolution, die ja selber nichts anderes ist als schrittweise Verfeinerung von biologischen Kreaturen, nicht alle Eigenschaften des Menschen erklärt, die anderer Tiere aber schon. Wenigstens wagt er es nicht, das Wort „Intelligent Design“ in den Mund zu nehmen. . .

Was sind nun die Eigenschaften, von denen Bringsjord meint, sie machten eine Person aus und könnten nur dem Menschen zugesprochen werden? Die Fähigkeit, in einer Sprache zu kommunizieren, „Freier Wille“, Kreativität, phänomenales Bewusstsein und robustes, abstraktes Denken. Das hört sich natürlich wunderbar an, doch wenn wir ein durchschnittliches Exemplar der Gattung *homo sapiens sapiens* heranziehen, sagen wir, den typischen Großstadtneurotiker, dann ist es fraglich, ob Bringsjord einen solchen als ultimative Referenz für „Fähigkeit, in einer Sprache zu kommunizieren“, „Freier Wille“, „Kreativität“ und insbesondere „robustes, abstraktes Denken“ anerkennen möchte. Über phänomenales Bewusstsein kann man sicher noch streiten.

Die von den Autoren vertretene Meinung, dass es klar scheine, dass „KI nur über genug Feuerkraft verfüge, Zombianimals bis zur Stufe eines Schimpansen hervorzubringen“, dass aber „Personen vom jetzigen Standpunkt der KI unerreichbar wären“, wirft die Frage auf, ob die Autoren wirklich nur zeigen wollen, dass Maschinen kein Bewusstsein haben können. Vielmehr impliziert ihre Meinung, dass auch (sogar höhere) Tiere keines haben! Mangels anderer Kandidaten für Bewusstsein können wir ihre Einstellung gleich auf die Aussage reduzieren: „Nur Menschen haben Bewusstsein.“. Es wäre interessant, ob die Autoren ihre Meinung auch dann noch aufrecht erhalten würden, wenn sie sich darüber mit Kollegen aus dem nächsten Institut für vergleichende Verhaltensforschung bei Primaten unterhielten. . .

6 Gegenargumente

In weiser Voraussicht, dass viele Gegenargumente gegen ihre Ansichten vorgebracht würden, haben sich die Autoren auch mit potentiellen solchen auseinandergesetzt. Da wird zuerst auf den berechtigten Einwand eingegangen, dass dieser Artikel im Prinzip nur auf dem Standpunkt beharrt, dass KI kein Be-

wusstsein hervorbringen könne. Nach abschweifenden Argumenten, dass dies ohnehin schon längst anderswo bewiesen (sic!) wäre, verteidigen sich Bringsjord et al. damit, dass sie sich hier nur mit KI im „wirklichen Leben“ mit „echten Praktikern“ auseinandersetzen. Und dass das Resultat ihrer Erkenntnisse sei, dass es ein klarer Fakt wäre, dass es nicht den geringsten Hinweis gäbe, wie man einem Roboter oder einem künstlichen Agenten soetwas wie phänomenales Bewusstsein einhauchen könne. Dass wir auch nicht wissen, wie wir selbiges verhindern können, scheint die Autoren nicht weiter zu stören, während es bei dem einen oder anderen Leser den Verdacht erwecken könnte, dass wir es hier mit einem völlig unsinnigen Konzept zutun haben. . .

Doch es wird noch besser: auf den nächsten Einwand, dass doch auch der Mensch durch Evolution aus primitiveren, offenbar „weniger bewussten“ Lebensformen hervorgegangen sei, dass also Bewusstsein inkrementell entstehen könne, antworten die Autoren kühn, um nicht zu sagen mit halsbrecherischem Wagemut, dass dieser Einwand bei weitem nicht fundiert wäre! Nämlich dass es noch immer ein offenes Problem sei, ob Evolution aufgrund eines sogenannten „Wallace’schen Paradoxons“ zurückgewiesen werden müsse! Möglicherweise ist es den Autoren entgangen, dass Evolution insofern ein Faktum ist, als Artenbildung sowohl in freier Wildbahn als auch im Labor beobachtet und aufgezeichnet wurde — bereits oftmals²! Die Evolutionstheorie bezieht sich darüber hinaus auf die Annahme, dass diese Prozesse in der Vergangenheit ebenso abgelaufen sind, wie sie es heute tun. Das kann zwar nicht bewiesen werden, da die Zeit nicht zurückgedreht werden kann, aber zahllose Fossilien- und Artefaktfunde unterstützen diese Hypothese sehr, sehr stark. Alfred Wallace, ein Zeitgenosse Darwins, der wohl heute aufgrund obiger Funde wahrscheinlich anders argumentieren würde, konnte sich damals nicht erklären, warum die Evolution Gehirne hervorgebracht hat, die Dinge vermögen, die für das Überleben nicht notwendig sind. Eigentlich sollte es nicht verwundern, dass notwendige Fähigkeiten auch unnötig bedingen können. So müssen Seehunde in freier Natur nicht mit Bällen jonglieren können, mit der dafür notwendigen Geschicklichkeit können sie aber gewiss etwas anfangen. Um diesen Punkt abzuschließen, die Autoren zeigen hier mit ihren Argumenten ein fantastisches Unverständnis grundlegender naturwissenschaftlicher Begriffe und Erkenntnisse und wären gut beraten, sich in diesem Bereich etwas eingehender zu informieren, bevor sie haltlose Behauptungen aufstellen.

Auf die nächsten Einwände, nämlich dass Tiere phänomenales Bewusstsein bräuchten, um X (eine beliebige Aktion) zu vollbringen, und darauf, dass kognitive Wissenschaften anders an das Problem herangingen als die KI, wollen wir hier nicht weiter eingehen. Zwar ist es vorstellbar, dass manche sich auf diese Diskussion einlassen würden, aber wenn man den Standpunkt vertritt, dass weder das Vorhandensein noch die Absenz von phänomenalem Bewusstsein überprüfbar sind, erscheint die Präsentation dieser Einwände und deren Abwehr durch Bringsjord et al. wie das Einschlagen auf Strohmänner.

Der vorletzte Einwand sagt (richtigerweise) aus, dass die Latte für Menschen heimlich angehoben wurde. Hier versteigen sich die Autoren in wüsten Spekulationen über des Menschen Fähigkeit, mit infinitären Logiken umgehen zu können. Da wird zwar richtig erwähnt, dass die gültigen Formeln einer bestimmten solchen Logik durch Turingmaschinen nicht aufgezählt werden können, doch

²Umfangreiche Informationsseite dazu: <http://www.talkorigins.org>

der Schluss, dass Menschen hier der Maschine überlegen wären, da sie sich mit formalen (sic!) Systemen beschäftigten, die Aussagen über solche Formeln machen, ist mehr als suspekt. Wir warten gespannt darauf, wie Bringsjord et al. uns demonstrieren, wie ihre überlegene menschliche Intelligenz es zu einer solchen Aufzählung bringt...

Schlussendlich muss auch noch Turing herhalten, indem ihm unlogische Argumentation untergeschoben wird, was unhaltbar ist, wenn man den Originaltext liest: das Argument bezieht sich auf die Fähigkeit von Maschinen, kreativ zu sein, also etwas Originelles hervorzubringen. Lady Lovelace, eine Mitarbeiterin von Charles Babbage, dem Erfinder der „Analytical Engine“, meinte, dass diese Fähigkeit von Computern nicht erbracht werden könne. Laut Bringsjord et al. widerspricht sich Turing, da er einerseits zustimmt, dass dies auf die erwähnte Maschine zutrifft, aber gleich darauf festhält, sie sei zu universellen Berechnungen fähig. Liest man den Originaltext von Turing, so wird allerdings klar, dass er dies so nicht gesagt hat! Vielmehr stellt er klar heraus, dass „Lady Lovelace zur damaligen Zeit nicht über Hinweise verfügte, ob die betreffende Maschine diese Eigenschaft hatte oder nicht“. Das ist ein ganz anderes Argument! Auch die weiteren Einwände Turings werden hoffnungslos verzerrt dargestellt: als ob Turing nur an Programmierfehler gedacht hätte, als er schrieb, dass ihn die Resultate von Computern manchmal überraschen! Sein Argument sagt aus, dass er manchmal über die Konsequenzen der Berechnung des Computers *nicht ausreichend nachgedacht* hat und deswegen überrascht wird. Dies schliesst vielleicht Programmierfehler ein, kann sich aber auch darauf beziehen, dass Turing nicht ausreichende Weitsicht hatte, um ein bestimmtes Ergebnis zu erwarten. Anders ausgedrückt: der Computer dachte schneller bzw. weiter und konnte daher etwas Unerwartetes finden.

7 Zusammenfassung: Turings Fehler

Wir wollen hier nur auf die allerletzte Aussage eingehen: dass Turing eines Tages als der Mensch bekannt sein würde, der die rechenbetonte Natur von Tieren erkannte, aber in religiösem Glauben ausdrückte, dass *Personen* (in der Definition von Bringsjord et al.) von Natur aus letztendes gleich wären. Nun denn, Turing vertritt zumindest konsistent eine Meinung, während dies offensichtlich den Autoren des behandelten Artikels schwer fällt: denn sonst würden sie nicht ständig hin- und herschwanken, ob nun Tiere „rechenbetonte Natur“ haben oder doch mehr sind (wozu dann „Zombanimals“, wenn normale Tiere auch kein Bewusstsein haben?) oder ob Menschen vielleicht doch Tiere sind, wie sie einmal meinen, oder doch nicht so ganz andernorts.

8 Anmerkung zur Analyse des Artikels

Sigmund Freud würde bei den Autoren dieses Artikels wahrscheinlich diagnostizieren, woran die Menschheit offenbar schon seit Ewigkeiten leidet: narzistische Störungen, durch die Menschen sich selbst betrügen, mehr zu sein, als sie sind. Entgegen allen Erkenntnissen der Wissenschaft führt der Mensch verbissene Rückzugsgefechte vom selbsterbauten Sockel als Krone der Schöpfung. Zu den narzistischen Kränkungen, die uns durch Galileo Galilei astronomisch

aus dem Mittelpunkt des Universums an einen unbedeutenden Ort katapultiert, durch Charles Darwins Einsichten Kenntnis über unsere evolutionäre Herkunft beschert, durch Sigmund Freud Einblick in unser dunkles Unterbewusstsein gewährt, durch Errungenschaften der Genetik unseren nicht so einzigartigen Platz unter den Lebewesen besiegelt hat, gesellt sich langsam aber unaufhaltsam die Einsicht, dass der Mensch auch intellektuell nicht notwendigerweise das Non-Plus-Ultra im Universum ist. Noch so viele Störversuche durch Verweis auf wissenschaftlich nicht sinnvoll überprüfbare und daher letztendes bedeutungslose Eigenschaften wie *phänomenales Bewusstsein* werden das nicht ändern.

Die Abkehr vom historisch gewachsenen, dualistischen Denken, wie es schon seit Menschengedenken unser Schicksal prägt und die Welt unbegründet in „bewusst“ und „unbewusst“, „lebendig“ und „unbelebt“, „Mensch“ und „Tier“, etc. trennt, ist immer noch eine Herausforderung unserer Zeit, wie dieser Artikel nur zu deutlich beweist. . .

Aktuelle naturwissenschaftliche Erkenntnisse aus Experimenten im neurophysiologischen aber auch im psychologischen Bereich zeigen, dass viele Annahmen, die wir mit dem Begriff *Bewusstsein* verbinden, bei näherer Betrachtung mit der Realität nicht im Einklang stehen. Manche Wissenschaftler, z.B. Susan Blackmore³, gehen den mutigen Schritt, die von uns üblicherweise verwendeten Definitionen von Bewusstsein anzuzweifeln, sie gar als „Illusion“ zu bezeichnen. Hoffentlich führt das dazu, dass Studenten in Hinkunft die Mühe erspart bleibt, solche Artikel wie die von Bringsjord et al. abzuhandeln. . .

Literatur

[BCN00] Selmer Bringsjord, Clarke Caporale, and Ron Noel. Animals, zombanimals, and the total turing test. *Journal of Logic, Language and Information*, 9(4):397–418, 2000.

³<http://www.susanblackmore.co.uk>