

Können Computer denken?

Eine Annäherung an die philosophische Schlüsselfrage der KI-Forschung

0. Inhaltsverzeichnis

0.	INHALTSVERZEICHNIS	2
1.	EINLEITUNG	3
2.	DIE TURINGMASCHINE	4
3.	DER TURINGTEST	8
4.	DAS ARGUMENT DES CHINESISCHEN ZIMMERS	10
5.	ERWIDERUNGEN AUF DAS ARGUMENT DES CHINESISCHEN ZIMMERS	14
6.	ZUSAMMENFASSUNG UND SCHLUSSFOLGERUNGEN	17
7.	ANHANG	19
A.	SPRACHANALYSE-PROGRAMME	19
I.	ELIZA	19
II.	PARRY	19
III.	HAM-ANS	20
IV.	FRUMP	20
V.	PAM	20
VI.	SHRDLU	20
VII.	A.L.I.C.E.	21
B.	BIBLIOGRAPHIE	21

1. Einleitung

Seit Jahrzehnten wird in der KI-Forschung (der Erforschung Künstlicher Intelligenz) die philosophische Schlüsselfrage „Können Computer denken?“ diskutiert. Diese Schlüsselfrage führt unmittelbar zu zwei Folgefragen, nämlich „Was ist ein Computer?“ und „Was ist Denken?“

Die erste Folgefrage mündet in technische Erklärungen der Funktionsweise von Computern und damit grundsätzlich in die Domänen der Mathematik und Physik. Auf dieser technischen Ebene der KI-Forschung dreht sich die Diskussion der philosophischen Schlüsselfrage um die Leistung von Computern: Dem technologischen Fortschritt werden mathematische Theorien als Beweis der eingeschränkten kognitiven Leistungsmöglichkeit oder logische Einwände als Beweis der eingeschränkten kognitiven Leistungsfähigkeit entgegengehalten.

Die zweite Folgefrage mündet in das Körper/Geist-Problem, das bereits seit der Antike untersucht und heute unter dem Namen *Bewusstseinsforschung* grundsätzlich als Domäne der Medizin erforscht wird. Auf dieser ‚geistigen‘ Ebene der KI-Forschung dreht sich die Diskussion der philosophischen Schlüsselfrage um die Ergründung von Bewusstsein, Intentionalität, Subjektivität und geistiger Verursachung (im Folgenden als *mentale Prozesse* bezeichnet): Dem Dualismus, der Körper und Geist beziehungsweise Hirn und Geist trennt und den Geist als zusätzliche Substanz betrachtet, wird der Materialismus entgegengehalten, der den Geist als Teil des Körpers beziehungsweise als Resultat der Hirntätigkeit versteht.

In der vorliegenden Arbeit möchte ich einen logischen Einwand auf der technischen Ebene diskutieren, nämlich das Argument des chinesischen Zimmers. Dazu wird zuerst exemplarisch die Funktionsweise von Computern erläutert, anschliessend wird mit dem Turingtest ein von vielen akzeptiertes Kriterium zur Bemessung ihrer kognitiven Leistungsfähigkeit vorgestellt. Dieses Kriterium wie auch die kognitive Leistungsfähigkeit allgemein werden danach unter Berücksichtigung der Erläuterungen zur Funktionsweise sowie des Arguments des chinesischen Zimmers eingehend untersucht; abschliessend werden einige Er widerungen auf die Erkenntnisse dieser Untersuchung angesprochen. Insgesamt sollte damit eine vorläufige Antwort auf die philosophische Schlüsselfrage der KI-Forschung gegeben werden können.

2. Die Turingmaschine

Bevor irgendeine Aussage über die kognitive Leistungsfähigkeit von Computern gemacht werden kann, muss ein ungefähres Verständnis ihrer Funktionsweise vorliegen. In diesem Abschnitt soll die Funktionsweise von Computern deshalb anhand eines einfachen Beispiels der Fragestellung dieser Arbeit angemessen definiert werden.¹

Nehmen wir einmal an, wir möchten schriftlich die Addition der Zahlen 123 und 90 durchführen, wofür wir über ein Blatt Papier und einen Stift verfügen. Wir notieren zuerst die beiden Zahlen, und zwar so, dass deren einzelnen Ziffern von hinten her untereinander zu liegen kommen. Dann beginnen wir, die hintersten Ziffern zusammenzuzählen: Wir erhalten 3, notieren 3 und merken uns den Übertrag 0. Dann rücken wir zu den Ziffern links davon und zählen diese sowie den Übertrag zusammen: Wir erhalten 11, notieren 1 und merken uns den Übertrag 1. Dann rücken wir erneut zu den Ziffern links davon und zählen diese sowie den Übertrag zusammen: Wir erhalten 2, notieren 2 und sind fertig; das Resultat lautet 213.

$$\begin{array}{r} 123 \\ + 90 \\ \hline 213 \end{array}$$

Diesem Additionsvorgang liegen zwei Voraussetzungen zugrunde: Die erste Voraussetzung ist das Wissen um Anzahl und Aussehen der zur Verfügung stehenden Ziffern, nämlich 1234567890. Die zweite Voraussetzung ist das schrittweise Durchlaufen von Handlungsanweisungen, die allgemein formuliert ungefähr wie folgt lauten:

1. Wenn noch nichts notiert wurde: Notiere die beiden Zahlen, so dass deren einzelnen Ziffern von hinten her untereinander zu liegen kommen und merke Übertrag = 0 sowie Position = 1.
2. Zähle die beiden Ziffern, die von hinten gezählt an der gemerkten Position stehen, sowie den gemerkten Übertrag gemäss folgender Tabelle zusammen:

¹ Vgl. Weizenbaum: Die Macht der Computer, S. 65-154.

Ziffer 1	Ziffer 2	gem. Übertrag	Resultat	neuer Übertrag
3	0	0	3	0
2	9	0	1	1
1	0	1	2	0
usw.				

3. Notiere das Resultat unter den beiden zusammengezählten Ziffern und merke den neuen Übertrag sowie Position = gemerkte Position + 1.
4. Wenn von hinten gezählt an der gemerkten Position keine weiteren Ziffern vorhanden sind: Notiere den gemerkten Übertrag vor den bisher notierten Resultatziffern und beende die Addition. Ansonsten beginne wieder von vorne.

Ein Schritt ist ein einmaliges Durchlaufen dieser Handlungsanweisungen, wobei jeder Schritt aus einem Ausgangszustand, einer Reihe von Handlungen und einem Endzustand besteht. Joseph Weizenbaum hat diese Handlungsanweisungen mit einem Spiel verglichen: Ein neuer Spieler kommt an die Reihe: Das Spiel weist eine bestimmte Spielsituation auf; der Spieler handelt gemäss den Spielregeln; das Spiel weist eine neue Spielsituation auf. Ein neuer Spieler kommt an die Reihe usw.

Diese ausführliche Beschreibung eines an sich sehr einfachen Additionsvorgangs ermöglicht einige wesentlichen Feststellungen:

- Eigentlich werden für den beschriebenen Additionsvorgang gar keine mathematischen Kenntnisse und damit auch kein Zahlenverständnis benötigt; wir brauchen beispielsweise nicht zu wissen, dass eins und eins zwei ergeben, dass drei auf zwei folgt oder dass wir ein dezimales Zahlensystem verwenden – das schrittweise Befolgen der Handlungsanweisungen führt uns zwangsläufig zum richtigen Resultat.
- Eigentlich ist die Verwendung von Ziffern für den beschriebenen Additionsvorgang unnötig und es kann auch mit beliebigen Symbolen operiert werden; wir können beispielsweise anstelle der zehn Ziffern zehn Buchstaben verwenden - wir brauchen nur die Handlungsanweisungen (und natürlich die erste Voraussetzung) entsprechend anzupassen und die nötigen Umwandlungen vornehmen.

- Eigentlich kann nicht nur mit beliebigen Symbolen, sondern auch mit einer beliebigen Anzahl Symbole operiert werden; wir können anstelle des dezimalen Zahlensystems beispielsweise ein hexadezimals oder ein binäres Zahlensystem verwenden und erhalten stets dasselbe Resultat - wir brauchen nur die Handlungsanweisungen (und natürlich die erste Voraussetzung) entsprechend anzupassen und die nötigen Umwandlungen vornehmen.
- Eigentlich können nicht nur mit einer beliebigen Anzahl beliebiger Symbole operiert, sondern sogar beliebige Operationen durchgeführt werden; wir können beispielsweise für die Addition ein anderes Vorgehen wählen oder anstelle der Addition eine Subtraktion durchführen – wir brauchen nur die Handlungsanweisungen entsprechend anzupassen.

Aus diesen wesentlichen Feststellungen folgt, dass sich mathematische Operationen als eine Reihe einfacher Handlungsanweisungen beschreiben lassen. Diese Erkenntnis nutzte Alan Turing 1936 zur ersten Definition der Funktionsweise von Computern.²

Turing definierte eine Maschine, die über ein Magnetband, also eine Reihe positiver und negativer Ladungen, sowie einen Magnetkopf verfügt, der diese Ladungen (aufgrund der magnetischen Kräfte) liest und entsprechend gewisser Handlungsanweisungen mit anderen Ladungen überschreibt, das heisst die Ladungen transformiert; dabei verändert das Verschieben des Magnetbands nach jeder Transformation die Position des Magnetkopfs. Entsprechend werden die oben aufgeführten Voraussetzungen so angepasst, dass erstens nur zwei ‚Ziffern‘ berücksichtigt werden, zweitens die ‚Zahl‘ in einer langen Reihe geschrieben (und damit immer nur eine ‚Ziffer‘ an einer bestimmten Position berücksichtigt) wird und drittens Handlungsanweisungen vorliegen, die jeweils Angaben über 1. den aktuellen Übertrag, 2. die aktuelle Ladung, 3. den neuen Übertrag, 4. die neue Ladung und 5. die Verschiebung der Position enthalten.

Die von Turing definierte Maschine war damit in der Lage, binäre Berechnungen durchzuführen. Auch wenn heutige Computer anders aufgebaut und deshalb massiv leistungsfähiger sind, funktionieren sie vom Prinzip her noch immer genau gleich: Sie transformieren gemäss Handlungsanweisungen schrittweise positive und negative Ladungen.

² Turing: On Computable Numbers.

Für die Fragestellung dieser Arbeit kann die Funktionsweise von Computern damit als *Konstruktion zum schrittweisen Durchlaufen vorgegebener Handlungsanweisungen* definiert werden.

3. Der Turingtest

Nachdem im vorderen Abschnitt die Funktionsweise von Computern erläutert worden ist, soll in diesem Abschnitt ein Kriterium zur Bemessung ihrer kognitiven Leistungsfähigkeit vorgestellt werden; sollte für die Bemessung der kognitiven Leistungsfähigkeit von Computern ein aussagekräftiges Kriterium vorliegen, würde dies den entscheidenden Schritt hin zur Beantwortung der philosophischen Schlüsselfrage der KI-Forschung ermöglichen. Ein solches von vielen Forschern akzeptiertes Kriterium wurde 1950 wiederum von Alan Turing formuliert.³

Turing suchte analog der vorliegenden Arbeit eine Antwort auf die Frage „Können Maschinen denken?“, doch er war der Ansicht, die Folgefrage „Was ist Denken?“ könne nicht beantwortet werden und entwickelte deshalb einen Test, um die kognitive Leistungsfähigkeit von Maschinen unter Ausklammerung der Bewusstseinsforschung zu bemessen. Dieser Test stützt sich auf ein Beispiel, das von zwei Testpersonen und einem Richter handelt. Die drei befinden sich in verschiedenen Räumen, wobei der Richter mit den Testpersonen über eine Konsole kommunizieren kann. Eine der beiden Testpersonen ist männlich, die andere weiblich, und der Richter muss mittels Kommunikation über die Konsole herausfinden, welche der beiden Testpersonen der Mann und welche die Frau ist. Dabei versucht eine der Testpersonen, dem Richter zu helfen, und die andere, den Richter zu täuschen. Der Richter und die helfende Testperson gewinnen, wenn der Richter Mann und Frau erkennt, die täuschende Testperson gewinnt, wenn der Richter Mann und Frau nicht erkennt. Um die kognitive Leistungsfähigkeit von Maschinen zu bemessen, ersetzte Turing die beiden Testpersonen des Beispiels durch einen Menschen und eine Maschine; die Maschine besteht den Test folglich dann, wenn der Richter Mensch und Maschine nicht erkennen kann.

Als Turing dieses Kriterium 1950 formulierte, waren Computer weit davon entfernt, seinen Test zu bestehen; eine gewisse Dialogfähigkeit wiesen sie erstmals um 1966 mit dem Sprachanalyse-Programm ELIZA⁴ auf. In der Folge wurde das Verhalten beispielsweise mit PARRY⁵, HAM-ANS⁶, FRUMP⁷, PAM⁸ und SHRDLU⁹ immer intelligenter, doch bis hin

³ Turing: Computing Machinery and Intelligence.

⁴ Vgl. S. 19.

⁵ Vgl. S. 19.

⁶ Vgl. S. 20.

zum heutigen Vorzeigebeispiel A.L.I.C.E.¹⁰ blieben gewisse Mängel bestehen: Folglich musste die seit 1991 von Hugh Loebner ausgesetzte Belohnung von US-\$ 100'000.- für das erste Sprachanalyse-Programm, welches den von Turing entwickelten Test vollumfänglich erfüllt, bisher nicht ausbezahlt werden¹¹. In der jährlich stattfindenden Bewertung erhielt A.L.I.C.E. im Jahr 2001 als beste Bewertung eines Computers immerhin 14 Punkte, während die schlechteste Bewertung eines Menschen 19 Punkte betrug - bedeutet dies also, dass 73% der Schwierigkeiten gelöst wurden, einen Computer zum Denken zu bringen? Oder bedeutet dies gar, dass Computer zu 73% denken können?

⁷ Vgl. S. 20.

⁸ Vgl. S. 20.

⁹ Vgl. S. 20.

¹⁰ Vgl. S. 21.

¹¹ <http://www.loebner.net/Prizef/loebner-prize.html> (April 2003).

4. Das Argument des chinesischen Zimmers

Der im vorderen Abschnitt vorgestellte Turingtest als Kriterium zur Bemessung der kognitiven Leistungsfähigkeit von Computern wie auch die kognitive Leistungsfähigkeit allgemein sollen nun unter Berücksichtigung der Erläuterungen zur Funktionsweise von Computern sowie des Arguments des chinesischen Zimmers eingehend untersucht werden.

Im vorderen Abschnitt wurde erzählt, dass Turing versuchte, die nach seiner Ansicht nicht zu beantwortende Frage „Was ist Denken?“ auszuklammern, indem er eine Umformulierung der Schlüsselfrage „Können Computer denken?“ in „Gibt es Maschinen, die als Dialogpartner für Menschen gehalten werden können?“ vornahm. Dabei stützte er sich implizit jedoch auf folgende Annahmen:

1. Denken bedeutet, dialogfähig zu sein.
2. Dialogfähig zu sein bedeutet, als Dialogpartner für einen Menschen gehalten zu werden.

Die Absicht von Turing, mittels Umformulierung die Frage „Was ist Denken?“ auszuklammern, scheitert also, weil eine Antwort auf diese Frage für die Umformulierung bereits vorausgesetzt wird. Und die von ihm vorausgesetzte Antwort ist zudem fragwürdig, denn Dialogfähigkeit impliziert nicht zwingend Denken: Wird eine Maschine als Dialogpartner für einen Menschen gehalten, heisst dies bloss, dass sie das Dialogverhalten eines Menschen einwandfrei simuliert hat und nicht, dass sie auch dieselben mentalen Prozesse durchlaufen hat.

Diesen Einwand hat John R. Searle 1980 in Form des Arguments des chinesischen Zimmers deutlich gemacht¹²: Dieses Argument geht von einer Person aus, die keinerlei Chinesisch versteht und auch mit den chinesischen Schriftzeichen nicht vertraut ist. Ausgestattet mit viel Notizpapier, einem Stift sowie einer umfangreichen Sammlung von Handlungsanweisungen wird sie allein in ein Zimmer gesperrt, das als einzige Verbindung zur Aussenwelt über zwei Schlitze verfügt; diese werden als Eingabeschlitz und Ausgabeschlitz bezeichnet und erlauben das Durchreichen von kleinen Zetteln. Durch den Eingabeschlitz erhält die Person nun einen

¹² Searle: Geist, Hirn und Wissenschaft, S. 30-36.

Zettel mit einer Frage in chinesischen Schriftzeichen, die sie natürlich nicht versteht. Sie bearbeitet diese Schriftzeichen deshalb unter Zuhilfenahme von Notizpapier und Stift gemäss den Handlungsanweisungen solange, bis sie eine Reihe anderer chinesischer Schriftzeichen, die sie natürlich ebenfalls nicht versteht, als Antwort zusammengetragen hat; dann zeichnet sie diese anderen Schriftzeichen auf die Rückseite des Zettels und steckt diesen in den Ausgabeschlitz. Damit kann die Person eine chinesische Frage auf Chinesisch beantworten, ohne ein Wort chinesisch zu verstehen - genau genommen kann die Person sogar auf eine Frage eine Antwort geben, ohne die Frage oder die Antwort zu verstehen. Natürlich würde ein solcher Vorgang eine enorme Anzahl von Handlungsanweisungen voraussetzen und entsprechend lange dauern; dennoch müsste der Person im Zimmer (beziehungsweise dem Zimmer mit der Person) dadurch Dialogfähigkeit attestiert werden. Doch zwischen der Person im Zimmer und einem Chinesen, die ja beide einen chinesischen Dialog führen können, besteht ein grosser Unterschied: Die Person im Zimmer kann nur Antworten geben, die sie mittels der vorgegebenen Handlungsanweisungen zusammentragen kann; ein Chinese hingegen kann auch auf unvorhergesehene Fragen angemessene Antworten geben.

Genau dieser Unterschied besteht auch zwischen Menschen und Computern: Menschen verfügen über intellektuelle Eigenschaften wie beispielsweise Bewusstsein und Vernunft, die es ihnen ermöglichen, einen Bezug zwischen dem Gedachten und seiner Bedeutung herzustellen und das Gedachte damit zu *verstehen*. Computer hingegen durchlaufen lediglich schrittweise die vorgegebenen Handlungsanweisungen, womit sie...

1. ...die Bedeutung der positiven und negativen Ladungen nicht erkennen können; Computer können also (in binäre Symbolketten umgewandelte) Zahlen, Texte, Bilder, Klänge und anderes nicht voneinander unterscheiden. Dies entspricht in etwa einem Funker im Militär, der eine chiffrierte Nachricht übermittelt: Es könnte sich dabei sowohl um einen platonischen Dialog wie auch um reinen Datenmüll handeln.
2. ...den Sinn der von ihnen vorgenommenen Transformationen nicht verstehen können; Computer können also nicht rechnen, sondern befolgen lediglich (aus physikalischen Gründen) Handlungsanweisungen, die auch eine Berechnung sein können. Dies entspricht in etwa einem Kind, welches eine Uhr erhält, mit der Aufforderung nach Hause zu kommen, wenn der kleine Zeiger nach unten und der grosse Zeiger nach oben zeigen: Das Kind kommt pünktlich nach Hause, obwohl es kein Verständnis von Zeit hat und folglich auch nicht weiss, dass sechs Uhr ist.

Searle hat diesen Unterschied zwischen Menschen und Computern mit den Begriffen *semantisches Verständnis* und *syntaktische Operationen* umschrieben. Damit stellt sich natürlich die Frage, weshalb wir Menschen zu semantischem Verständnis in der Lage sein sollen, weshalb wir also über intellektuelle Eigenschaften wie beispielsweise Bewusstsein und Vernunft verfügen sollen und Computer nicht. Vermutlich sind diese Eigenschaften das Resultat unseres Hirns, unserer Sinnesorgane, ja unseres ganzen Körpers, also nicht konkrete, greifbare Bereiche im Sinne eines Homunculus, sondern systembedingte Konsequenzen der komplexen menschlichen Physik. Damit ist denkbar, dass künstliche Systeme (im Optimalfall) menschliches Verhalten funktional äquivalent simulieren können, ohne dabei jedoch zwingend dieselben intellektuellen Eigenschaften aufweisen zu müssen. Searle hat diese Theorie mit den Worten zusammengefasst:

„Es ist für den Geist nicht irrelevant, in menschlichen Hirnen realisiert zu sein.“¹³

- Der von Turing entwickelte Test, der korrekt umformuliert mit der Frage „Gibt es Maschinen, die das Dialogverhalten von Menschen einwandfrei simulieren können?“ zusammengefasst werden kann, kann somit als interessante Herausforderung für den technologischen Fortschritt, aber keinesfalls als Kriterium zur Bemessung der kognitiven Leistungsfähigkeit von Computern betrachtet werden. Damit liegt kein aussagekräftiges Kriterium vor, welches den entscheidenden Schritt in der Beantwortung der philosophischen Schlüsselfrage der KI-Forschung ermöglichen würde.
- Bei der Bemessung der kognitiven Leistungsfähigkeit von Computern allgemein müsste zudem zwischen *simuliertem Denken* (syntaktische Operationen ohne semantisches Verständnis) und *wirklichem Denken* (syntaktische Operationen mit semantischem Verständnis) unterschieden werden: Das simulierte Denken entwickelt sich abhängig vom technologischen Fortschritt und dürfte wohl zu einem immer intelligenteren Verhalten von Computern führen. Das wirkliche Denken hingegen, also das

¹³ Searle: Geist, Hirn und Wissenschaft, S. 39.

Durchlaufen mentaler Prozesse, wie sie auch Menschen durchlaufen, wird mit dem Argument des chinesischen Zimmers unabhängig vom technologischen Fortschritt grundsätzlich in Frage gestellt. Mit dieser Unterscheidung kann nun auch die am Ende des letzten Abschnitts aufgeworfene Frage beantwortet werden: Die Bewertung bedeutet demnach, dass Computer das menschliche Dialogverhalten zu 73% simulieren können.

5. Erwidernngen auf das Argument des chinesischen Zimmers

Obwohl Turings Test, wie im vorderen Abschnitt gezeigt, kaum als aussagekräftiges Kriterium zur Bemessung der kognitiven Leistungsfähigkeit von Computern betrachtet werden kann und mit dem Argument des chinesischen Zimmers wirkliches Denken bei Computern grundsätzlich in Frage gestellt wird, gibt es in der KI-Forschung bis heute viele Spezialisten, die Computern wirkliches Denken einräumen und dafür sogar an dem von Turing entwickelten Test als aussagekräftiges Kriterium festhalten. In diesem Abschnitt sollen kurz zusammengefasst die wichtigsten Erwidernngen dieser Spezialisten auf die bisherigen Erkenntnisse angesprochen werden.

Die wichtigsten Erwidernngen berufen sich allesamt auf die materialistische Theorie, wonach mentale Prozesse bloss das Resultat der Hirntätigkeit seien¹⁴. Zur Veranschaulichung wird auf die Anhänger des Vitalismus verwiesen, die zu Beginn des Zwanzigsten Jahrhunderts nach einer Substanz gesucht hätten, welche einem Körper das Leben ermögliche; erst später sei akzeptiert worden, dass die Funktionen der Zellen, der Organe, ja des Körpers insgesamt für das Leben verantwortlich seien. Analog werde jetzt nach einer Substanz gesucht, welche einem Körper mentale Prozesse ermögliche; doch auch hier müsse akzeptiert werden, dass die Funktionen der Hirnzellen, des Hirns, ja des Körpers insgesamt für die mentalen Prozesse verantwortlich seien.

Folglich behauptet beispielsweise Daniel C. Dennett¹⁵, das Argument des chinesischen Zimmers sei missraten, denn da im menschlichen Hirn auch nicht bloss eine Hirnzelle oder Hirnregion, sondern nur das Zusammenspiel aller Hirnzellen und Hirnregionen die mentalen Prozesse hervorrufe, dürfe beim Argument des chinesischen Zimmers auch nicht bloss die Person, sondern nur das Zusammenspiel von Person, Notizpapier, Stift und Regelwerk auf mentale Prozesse untersucht werden; es gebe kein Denkzentrum im Hirn, weshalb die Person auch nicht als Denkzentrum des chinesischen Zimmers betrachtet werden dürfe. Diese Unterscheidung sei wesentlich, weil beim Bearbeiten der chinesischen Schriftzeichen Tausende, ja Millionen von Regeln abgearbeitet und entsprechende Mengen Notizpapier verbraucht würden – und genau diese nur als Hilfsmittel eingestufteten Regeln und Notizen

¹⁴ Vgl. Dennett: Consciousness Explained.

¹⁵ Dennett: Consciousness Explained, S. 431-440.

entsprechen den Vorgängen und Status im menschlichen Hirn. Analog würden also die magnetischen Operationen von Computern den neuronalen Operationen von Hirnen entsprechen, womit diese Erwiderung wie folgt strukturiert werden kann:

1. Wenn eine Maschine ein bestimmtes menschliches Verhalten einwandfrei simulieren kann, dann müssen für diese Simulation eine vergleichbare Anzahl von Regeln abgearbeitet werden, die unbewusst auch das menschliche Hirn abarbeiten muss.
2. Mentale Prozesse und damit semantisches Verständnis sind einzig und zwingend Resultat syntaktischer Operationen.
3. Also durchlaufen Mensch und Maschine bei demselben Verhalten auch dieselben mentalen Prozesse; und weil Dialogfähigkeit eine Komponente des Denkens repräsentiert, kann das Dialogverhalten zur Bemessung der kognitiven Leistungsfähigkeit herangezogen werden.

Allerdings scheint der Zusatz *zwingend* in der zweiten Prämisse fragwürdig zu sein, denn auch wenn eingeräumt wird, dass die Person im chinesischen Zimmer nicht als Denkzentrum betrachtet werden darf, bleibt zweifelhaft, ob dem menschlichen Hirn funktional äquivalente künstliche Systeme auch zwingend über dieselben intellektuellen Eigenschaften verfügen.

Auch David J. Chalmers hält dem Argument des chinesischen Zimmers die eben angesprochene Erwiderung entgegen, ergänzt diese aber mit der Behauptung, das Argument vermische mit syntaktischen Operationen und semantischem Verständnis Ursache und Wirkung¹⁶: Das Abarbeiten von Regeln und mentale Prozesse seien nicht bloss zwei individuelle Eigenschaften, sondern das Abarbeiten von Regeln führe zu mentalen Prozessen. Entsprechend karikiert Chalmers das Argument des chinesischen Zimmers wie folgt:

¹⁶ Chalmers: *The Conscious Mind*, S. 313-332.

1. Computerprogramme sind abzuarbeitende Regeln.
2. Das Abarbeiten von Regeln genügt nicht, um semantisches Verständnis herzustellen.
3. Mentale Prozesse implizieren semantisches Verständnis.
4. Also genügen Computerprogramme nicht, um mentale Prozesse herzustellen.

1. Kuchenrezepte sind abzuarbeitende Regeln.
2. Das Abarbeiten von Regeln genügt nicht, um Geschmack herzustellen.
3. Kuchen implizieren Geschmack.
4. Also genügen Kuchenrezepte nicht, um Kuchen herzustellen.

In seiner Erwiderung scheint Chalmers aber ebenfalls zwei Dinge zu vermischen, und zwar Computer und Computerprogramme: Das Argument des chinesischen Zimmers bezieht sich nämlich nicht wie karikiert auf Computerprogramme, sondern auf Computer. Dies scheint insofern von Bedeutung, als bei Berufung auf die materialistische Theorie immer wieder die Komplexität der syntaktischen Operationen hervorgehoben wird, welche die Grundlage für das semantische Verständnis und damit die mentalen Prozesse sein soll. Damit scheint die Fähigkeit *Denken* aber am Inhalt des Gedachten gemessen zu werden, was fragwürdig wäre; denn das Berechnen von $123 + 90$ setzt verglichen mit $1 + 1$ wegen der höheren Komplexität wohl kaum ein höheres Bewusstsein voraus.

Doch diese Erwiderungen mit der materialistischen Theorie als Teil der Bewusstseinsforschung führen stark auf die ‚geistige‘ Ebene der KI-Forschung. Damit wird klar, dass das Argument des chinesischen Zimmers nur mit weiteren medizinischen Erkenntnissen betreffend unserer Hirnfunktionen in Verlegenheit gebracht werden kann.

6. Zusammenfassung und Schlussfolgerungen

Die vorliegende Arbeit soll eine Annäherung an die philosophische Schlüsselfrage der KI-Forschung „Können Computer denken?“ darstellen. Um dieses Ziel zu erreichen, sind die folgenden Diskussionspunkte aufgegriffen worden:

- Nach der Einleitung ist im zweiten Abschnitt am Beispiel eines einfachen Additionsvorgangs die Funktionsweise von Computern erläutert und eine der Fragestellung der Arbeit angemessene Definition eingeführt worden; demnach sind Computer *Konstruktionen zum schrittweisen Durchlaufen vorgegebener Handlungsanweisungen*.
- Im dritten Abschnitt ist mit dem Turingtest ein Kriterium zur Bemessung der kognitiven Leistungsfähigkeit von Computern vorgestellt worden, das sich auf Dialogfähigkeit konzentriert; eine anschließende Wertung des aktuellen technologischen Fortschritts hat gezeigt, dass Computer dieses Kriterium noch immer nicht erfüllen können.
- Im vierten Abschnitt ist das Kriterium wie auch die kognitive Leistungsfähigkeit allgemein unter Berücksichtigung der Erläuterungen zur Funktionsweise sowie des Arguments des chinesischen Zimmers eingehend untersucht worden: Das Kriterium scheidet nicht nur hinsichtlich der beabsichtigten Ausklammerung der Bewusstseinsforschung, sondern reduziert Denken fälschlicherweise auf Dialogfähigkeit. Das Argument des chinesischen Zimmers demonstriert, dass Computer auf vorgegebene Handlungsanleitungen beschränkt sind, während Menschen über intellektuelle Eigenschaften wie Bewusstsein oder Vernunft verfügen. Daraus folgt, dass 1. Computer den Inhalt der positiven und negativen Ladungen nicht erkennen können und 2. Computer den Sinn der von ihnen vorgenommenen Transformationen nicht verstehen können. Also ist denkbar, dass künstliche Systeme menschliches Verhalten funktional äquivalent simulieren können, ohne dabei zwingend dieselben intellektuellen Eigenschaften aufweisen zu müssen. Der Turingtest ist folglich als Kriterium zur Bemessung der kognitiven Leistungsfähigkeit von Computern verworfen und eine Unterscheidung in *simuliertes Denken* und *wirkliches Denken* eingeführt worden, wobei wirkliches Denken grundsätzlich in Frage gestellt worden ist.

- Im fünften Abschnitt sind kurz zusammengefasst die wichtigsten Er widerungen auf die vorhergehenden Erkenntnisse angesprochen worden, die sich allesamt auf die materialistische Theorie berufen: Demnach versage das Argument des chinesischen Zimmers einerseits, weil die Person als Denkzentrum betrachtet werde, jedoch analog der Hirntätigkeit das Zusammenspiel von Person, Notizpapier, Stift und Regelwerk die mentalen Prozesse hervorrufe, andererseits, weil das Argument Ursache und Wirkung vermische.

Mit diesem Vorgehen wurde versucht, den gewonnenen Eindruck zu vermitteln, dass sich die Diskussion um die philosophische Schlüsselfrage der KI-Forschung in letzter Zeit etwas totgelaufen hat: Während die mathematischen Theorien als Beweis der eingeschränkten kognitiven Leistungsmöglichkeit vom technologischen Fortschritt überrollt worden sind, wirken die Er widerungen auf die logischen Einwände als Beweis der eingeschränkten kognitiven Leistungsfähigkeit (und damit insbesondere auf das Argument des chinesischen Zimmers) eher fragwürdig. Dies liegt in erster Linie daran, dass die Bewusstseinsforschung trotz grosser medizinischer Fortschritte noch immer kein aussagekräftiges Kriterium zur Bemessung der kognitiven Leistungsfähigkeit hat formulieren und damit der KI-Forschung noch kaum Impulse hat verleihen können - was angesichts der zwar plausiblen, aber schwer fassbaren materialistischen Theorie auch nicht erstaunt. Letztlich dürfte eine Wiederbelebung der Diskussion aber nur mit wichtigen neuen Erkenntnissen der Bewusstseinsforschung möglich sein.

7. Anhang

A. Sprachanalyse-Programme

I. ELIZA¹⁷

...gilt als das erste Sprachanalyse-Programm und wurde um 1966 von Joseph Weizenbaum entwickelt. Ursprünglich war dieses Programm auf Magnetbändern implementiert und kommunizierte über eine Schreibmaschine. Ein erstes Magnetband nahm die Sprachanalyse der Eingaben vor, ein zweites Magnetband generierte anhand eines Skripts die Antworten. Das Skript konnte angepasst und so das *Verhalten* von ELIZA verändert werden.

Das Programm ist simpel aufgebaut, da es lediglich die Eingaben nach Schlüsselbegriffen durchsucht und Teile davon in der Antwort wieder aufführt. So antwortet ELIZA etwa auf die Eingabe „Ich hasse Menschen.“ mit „Warum hassen Sie Menschen?“ oder „Sind Sie sicher, dass Sie Menschen hassen?“. Entsprechend einfach gestaltet sich deshalb der Gesprächsverlauf, ziemlich oft kommt es zudem zu unsinnigen Antworten. Trotzdem kann mit ELIZA bereits eine Art Gespräch geführt werden.

II. PARRY¹⁸

...simuliert einen jungen Paranoiker, der im Rahmen einer Psychoanalyse seine aufgrund einer bestimmten *Vergangenheit* entstandenen Probleme vorträgt. Im Gegensatz zu ELIZA werden von PARRY 1. aufgrund der *Vergangenheit* Antworten gegeben, die über die getätigten Eingaben hinausgehen, und 2. zusätzlich zur Sprachanalyse über den gesamten Gesprächsverlauf verschiedene *Empfindungen* wie Angst, Ärger oder Vertrauen berechnet, womit das Programm für eine Änderung des *Verhaltens* keiner Skriptanpassung bedarf, sondern dieses aufgrund seiner *Empfindungen* selber verändern kann. Dies führt (nicht zuletzt dank der vorausgesetzten Paranoia) im Rahmen der Psychoanalyse zu einem eigentlichen Gesprächsverlauf mit wenigen unsinnigen Antworten.

¹⁷ Weizenbaum: Die Macht der Computer, S. 14-21.
<http://www.uwec.edu/jerzdg/orr/articles/IF/canon/Eliza.htm> (April 2003).

¹⁸ D'Avis: Können Computer denken?, S. 31-36.

III. HAM-ANS¹⁹

...erledigt im Auftrag eines Hotels telefonische Anfragen und Reservationen. Im Gegensatz zu ELIZA kann HAM-ANS 1. also *Hören* und *Sprechen*, 2. aufgrund einer *Hotelausbildung* zu logischen Annahmen und Schlussfolgerungen gelangen, die über die getätigten Eingaben hinausgehen, 3. sich an den gesamten Gesprächsverlauf *erinnern* und 4. dank einem ausgefeilten *Sprachverständnis* mit Füllworten, Andeutungen und sogar grammatikalisch inkorrekten Sätzen umgehen. Dies führt zu einem nahezu fehlerfreien Gesprächsverlauf im Rahmen der Anfragen und Reservationen eines Hotels.

IV. FRUMP²⁰

...gilt als das erste Sprachanalyse-Programm, welches nicht mehr die Grammatik, sondern die Bedeutung der Eingaben untersucht hat. Das Programm ist in der Lage, Texte wie beispielsweise Zeitungsberichte je nach Wunsch mehr oder weniger stark gekürzt zusammenzufassen. Im Gegensatz zu ELIZA kann FRUMP also einen längeren Text analysieren, strukturieren und die Bedeutung der einzelnen Sätze berechnen.

V. PAM²¹

...kann die kausalen Zusammenhänge zwischen eingegebenen einfachen Sätzen herstellen. Im Gegensatz zu ELIZA kann PAM also 1. ein gewisses *Textverständnis* vorweisen und 2. Antworten geben, die allein aufgrund des *Textverständnisses* über die getätigten Eingaben hinausgehen.

VI. SHRDLU²²

... wurde konzipiert, um analog einem Kleinkind mit Klötzchen zu spielen. Im Gegensatz zu ELIZA weist SHRDLU deshalb 1. ein beachtliches Wissen über die Klötzchenwelt und 2. Kamera und Greifarm auf, womit die Umgebung *gesehen* und Aktionen ausgeführt werden

¹⁹ D'Avis: Können Computer denken?, S. 64-69.

²⁰ D'Avis: Können Computer denken?, S. 41.
<http://newsblaster.cs.columbia.edu> (April 2003).

²¹ D'Avis: Können Computer denken?, S. 45-48.

²² D'Avis: Können Computer denken?, S. 61-64.

können. Dazu werden Anweisungen entgegengenommen und Fragen gestellt, beispielsweise wird auf die Eingabe „Ergreife die Pyramide!“ die Pyramide ergriffen oder „Welche Pyramide?“ nachgefragt, falls nicht ersichtlich sein sollte, welche Pyramide gemeint ist.

VII.A.L.I.C.E.²³

...ist eines der neueren Sprachanalyse-Programme und wurde als virtueller Chat-Partner entwickelt. Das Programm nimmt ähnlich wie ELIZA lediglich eine Sprachanalyse der Eingaben vor, erlaubt aber einen eigentlichen Gesprächsverlauf mit wenigen unsinnigen Antworten über beliebige Themen.

B. Bibliographie

Chalmers, David J.: **The Conscious Mind** – In Search Of A Fundamental Theory, New York 1996.

D’Avis, Winfried: **Können Computer denken?** – Eine bedeutungs- und zeittheoretische Analyse von KI-Maschinen, Frankfurt am Main / New York 1994.

Dennett, Daniel C.: **Consciousness Explained**, Boston / Toronto / London 1991.

Searle, John R.: **Geist, Hirn und Wissenschaft** – Die Reith Lectures 1984, Frankfurt am Main 1986.

Turing, Alan M.: **On Computable Numbers With An Application To The Entscheidungsproblem** in: Proc. London Math. Soc. Ser. 2-42 (1936), S. 230-265, <http://www.abelard.org/turpap2/tp2-ie.asp> (April 2003).

Turing, Alan M.: **Computing Machinery And Intelligence** in: Mind 59/236 (1950), S. 433-460, <http://www.abelard.org/turpap/turpap.htm> (April 2003).

Weizenbaum, Joseph: **Die Macht der Computer und die Ohnmacht der Vernunft**, Frankfurt am Main 1977.

²³ <http://www.alicebot.org> (April 2003).