

Leopold Franzens Universität

Institut für Psychologie

LV-Titel: Forschungsseminar: Neuere psychologische Fachliteratur (Gruppe D)

LV-Nummer: SS11. 720584

LV-Leiter: Univ-Prof. Dr. Leidlmair Karl

„The Symbol Grounding Problem“



„Nihil est in intellectu, quod non antea
fuerit in sensu“

John Locke (1632 - 1704), englischer
Philosoph und Aufklärer

Katharina Amann

0718485

C298

Katharina.amann@student.uibk.ac.at

Christian Klöble

0717546

C298

Christian.Kloeble@student.uibk.ac.at

Innsbruck, am 11.05. 2011

Inhaltsverzeichnis:

1. Einleitung.....	3
2. Geschichtliche Annäherung: Vom Behaviorismus zum Kognitivismus....	4
3. Symbolsysteme.....	6
4. Konnektionistische Systeme.....	9
5. The Symbol Grounding Problem.....	10
5.1 Turing-Test.....	10
5.2 Das chinesische Zimmer.....	11
5.3 Chinese Dictionary-go-round.....	12
6. Die menschlichen Fähigkeiten.....	14
6.1 Ikonische und kategoriale Repräsentationen.....	15
6.2 Symbolische Repräsentationen.....	15
6.3 Die komplementäre Rolle des Konnektionismus.....	17
7. Das hybride System.....	17
8. Kann künstliche Intelligenz das Symbol Grounding Problem überwinden?.....	18
Literaturverzeichnis.....	20

1. Einleitung:

Die Kognitionswissenschaft (engl. Cognitive Science) hat das Ziel kognitive Fähigkeiten, wie beispielsweise Wahrnehmung, Denken oder Lernen zu erfassen. Mittlerweile werden allerdings u.a. auch analoges und metaphorisches Denken sowie das Bewusstsein dazu gezählt. Als wissenschaftliche Disziplin des 21. Jahrhunderts beschäftigt sich die Cognitive Science nicht nur auf experimenteller sondern auch auf theoretischer Weise mit dem wissenschaftlichen Studium von Gehirn und Geist. (vgl. Schulke 2010)

Diese Arbeit entstand im Rahmen des Forschungsseminars „Neuere psychologische Fachliteratur“ vom Sommersemester 2011 an der Universität Innsbruck. Nach einer kurzen aber prägnanten Einführung in die Cognitive Science wurden Kleingruppen gebildet, die sich jeweils mit diversen Themengebieten auseinandersetzen sollten. Unsere Aufgabe war es das Problem der Cognitive Science, die eine Antwort auf das Gedankenexperiment des chinesischen Zimmers von John Searle darstellte, genauer zu betrachten. Dabei berufen wir uns in erster Linie auf den Text „The Symbol Grounding Problem“ von Steven Harnad und gehen hier anhand des Wissenschaftlers gestellten Fragen ein. Ebenso werden der geschichtliche Hintergrund sowie Symbolsysteme erläutert; weiteres wird auf das chinesische Zimmer sowie den Turing-Test eingegangen, ehe Lösungsvorschläge im Hinblick auf das Problem der Cognitive Science die Arbeit abrunden.

2. Geschichtliche Annäherung: Vom Behaviorismus zum Kognitivismus

Lange Zeit war es einzig auf behavioristischer Grundlage möglich, dem Problem der Psychologie auf empirischem Wege entgegenzutreten. Die Bezeichnung Behaviorismus, welche sich vom lateinischen Wort *behavior* ableitet, wurde Ende des 19. Jahrhunderts als psychologische Erklärung begründet (vgl. Lexikon für Psychologie und Pädagogik); deren beide Hauptvertreter Pawlow und Skinner sind Jedermann/Jederfrau bekannt, da man sich zu Anfängen des Psychologiestudiums bereits sehr intensiv mit ihnen auseinandergesetzt hat.

Nach Auffassung des Behaviorismus war es nicht erlaubt aus Gedanken bzw. etwaigen Vorgängen im Gehirn - auch nicht wenn ein bestimmtes Verhalten gezeigt wurde - Theorien abzuleiten. Einzig das Beobachtbare, welches selbsterklärend ist, war Gegenstand des Behaviorismus. Erst durch die schrittweise Einführung des Kognitivismus konnten akzeptable Schlüsse über die dem Verhalten zugrundeliegenden beobachtbaren Schlüsse gezogen werden.

Bezieht man sich wiederum auf den Begriff des Behaviorismus lässt sich erkennen, dass sich darunter alle Forschungsprogramme finden lassen, die der naturwissenschaftlichen, experimentellen bzw. objektiven Methode zugeordnet werden. Somit erscheint es plausibel, dass sich der Behaviorismus nicht bzw. nur kaum mit der Problematik des „Symbol Grounding Problems“ beschäftigt hat.

Seit der kognitiven Wende – geprägt durch die Anfang der 1970er Jahren zustande gekommenen starken Einbezügen von Kognitionen und Denkvorgängen in Theoriebildung und Forschung (vgl. Internet 1) - , welche einen Übergang vom Behaviorismus zum Kognitivismus darstellte, wurden auch unbeobachtbare Prozesse zu Untersuchungen herangezogen; des weiteren wurden subjektive Interpretationen von der Wissenschaft akzeptiert.

Die Kognitionswissenschaft zog eine Wirkungsweise mit sich, die weit über die Psychologie hinaus ging. So lassen sich heutzutage u.a. in der Linguistik, der Mathematik, der Philosophie aber auch der Soziologie oder der Hirnforschung Kognitionswissenschaftler finden.

Eine zentrale Bedeutung kommt der uns heute bekannten semantischen Interpretierbarkeit zu, welche vom theoretischen Vokabular des Kognitivismus bestimmt wird und zugleich eines der zentralsten sowie wichtigsten Gebiete in diesem darstellt. (vgl. Harnad 1990, S.335)

„Unter Miteinbezug verschiedenster Elemente sowohl aus der Philosophie des Geistes als auch aus der Kognitionswissenschaft brachte Fodor eine umfassende Theorie des Geistes hervor, welche von ihm selbst als „repräsentational“ bezeichnet wurde. Diese „repräsentationale“ Theorie „lässt sich am besten als Versuch begreifen, eine Antwort auf die Frage zu finden, wie mentale Zustände psychisch realisiert sein können.“ (vgl. Beckermann 2008, S.301) Fodor geht dabei von einem Drei-Ebenen-Modell aus:

1. Die oberste Ebene repräsentiert die intentionalen Zustände des Systems und der Kausalbeziehung, die innerhalb dieser Zustände herrscht.
2. Die zweite Ebene legt ihren Fokus auf die Realisierung der vorhandenen Kausalbeziehungen mittels Symbolverarbeitungsprozessen.
3. Auf der dritten Ebene findet die physische Realisierung der Symbolverarbeitungsprozesse der zweiten Ebene statt; sie wird auch als Implementationsebene bezeichnet.

Hauptproblem in Fodors Theorie bildet die Frage inwieweit intentionale Zustände sowie die zwischen ihnen bestehenden Kausalbeziehungen durch Symbolverarbeitungsprozesse realisiert werden können.

Dadurch, dass dieses 3-Ebenen Modell nicht nur in der Kognitionswissenschaft, sondern auch in der Computerwissenschaft sehr stark verbreitet ist (vgl. Beckermann 2008, S.301), stellt Fodors repräsentationale Theorie eine Analogie zum Computer dar. Dabei spricht er von einer in ihrer „Funktion“ übereinstimmenden Ähnlichkeit zwischen einem PC, welcher sowohl aus einer Hardware als auch aus einer Software besteht und dem eben genannten Verhältnis von Gehirn und Geist. Aufgrund dessen, dass die Software ontologisch abhängig ist, ist sie in einem gewissen Sinne auch wieder unabhängig und zwar so, dass sie genau beschrieben werden kann, ohne die Implantierung zu kennen. Setzt man dies in Beziehung zum Geist, lässt sich erkennen, dass durch die Kognitionswissenschaft eine Beschreibung auf abstrakter Ebene möglich ist, ohne dass dabei eine Beschreibung des Gehirns von Nöten ist.

Des Weiteren geht Fodor in seiner Theorie davon aus, dass es so etwas wie eine Sprache des Geistes gibt (language of thought), in der der Geist mit mentalen Repräsentationen arbeitet, die nach einem mentalen Syntax zu Gedanken zusammengesetzt werden. Seinem Werk „the language of thought“, welches 1975 publiziert worden ist, wurde eine fundamentale

Bedeutung zugeschrieben und lange Zeit galt dies sogar als eines der wichtigsten Werke der Kognitionstheorie. (vgl. Harnad 1990, S. 335)

Fodors Theorie wurde jedoch nicht von Kritik verschont, besonders in den letzten Jahrzehnten wurde der Vergleich mit dem PC (Hardware/Software) erheblich bemängelt. Begründet wurde diese Kritik dadurch, dass von Seiten des Konnektionismus ein realistischeres Modell des Geistes entwickelt wurde, welches auf die Trennung von Hardware und Software ganzheitlich verzichtet. Dabei können künstliche neuronale Netze kognitive Fähigkeiten simulieren, ohne dass sie überhaupt einen Syntax bzw. eine explizite Repräsentation besitzen. Nach Ansicht Fodors können solch charakteristische menschliche Fähigkeiten grundsätzlich nicht simuliert werden. (vgl. Rickheit, Hermann & Deutsch 2003, S.5)

Überdies geht Fodor davon aus, dass der Verstand ein Symbolsystem ist und die Kognition bzw. Erkenntnis stellt eine Symbolhandhabung dar. Die Möglichkeit komplexe Verhaltensweisen zu erzeugen wurde anhand der Symbolhandhabung auf empirischem Wege mittels Erfolgen auf dem Gebiet der künstlichen Intelligenz dargelegt. (vgl. Harnad 1990, S.335)

3. Symbolsysteme:

Um dem Terminus Symbolsysteme auf den Grund zu gehen, muss man sich erstmals intensiver mit den beiden Begriffen Symbol und System auseinandersetzen.

Der Begriff Symbol kommt aus dem griechischen und bezeichnet ein „Zeichen bzw. Sinnbild, das stellvertretend für etwas nicht Wahrnehmbares steht“ (Brockhaus 2002, Band 3 S. 338).

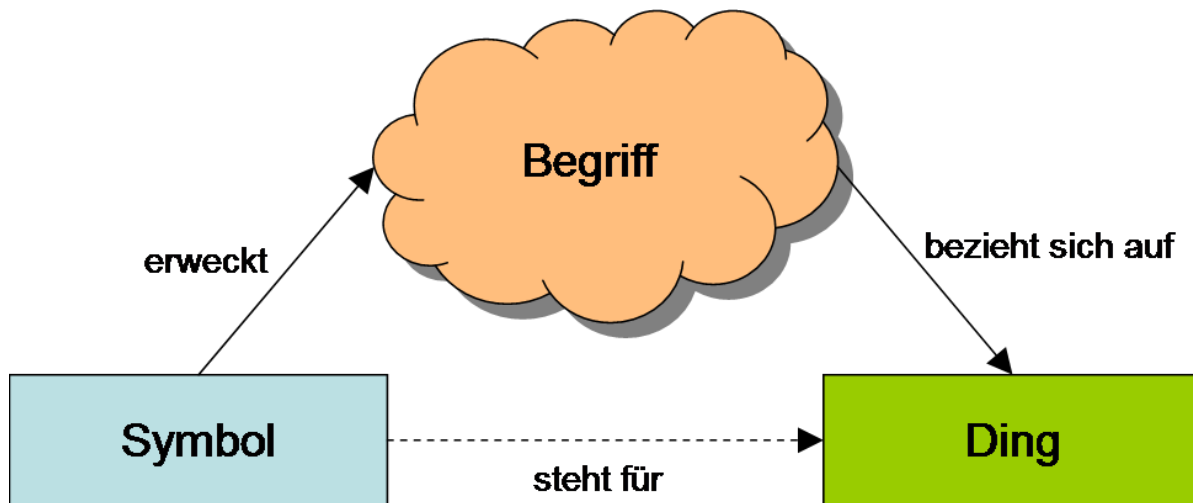


Abb.3: *semiotisches Dreieck*: Mittels dieser Graphik wird veranschaulicht, dass ein Zeichenträger sich nicht unmittelbar und direkt auf ein Ding bezieht, sondern dies durch einen mittelbaren Bezug aufgrund der Vermittlung/Vorstellung eines Begriffes geschieht. (vgl. Wörterbuch und Enzyklopädien auf der Akademik)

Dadurch, dass ein Symbol nicht immer offenkundig ist, gibt es etliche, die mehrere Bedeutungen beinhalten. Diese enorme Vieldeutigkeit erscheint oftmals sehr verschwommen oder auch nur vage interpretierbar. Der dahinterstehende Sinn eines Symbols deckt sich meist lediglich mit dem in Verbindung stehenden Kontext auf. Es gibt aber auch Symbole, welche künstlich geschaffen werden, so etwa die Zeichen am Flughafen, die uns sagen, an welchem Ort wir einchecken müssen bzw. wo es wieder zum Ausgang geht oder darauf aufmerksam gemacht wird, dass hier Rauchverbot herrscht, etc. Diese hier auftretenden Symbole erscheinen oftmals sehr einsichtig und dienen in weiterer Hinsicht zur Verständigung über die Sprachgrenzen hinaus. Andererseits existieren auch Zeichen, die erlernt werden müssen, wie beispielsweise die Schrift oder die Zahlzeichen. (vgl. Internet 2) Probleme lassen sich hingegen bei Abstraktionen, z.B. Liebe oder kulturellen Symbolen, wie beispielsweise Weihwasser finden, da ein jeder/ein jede darunter etwas anderes versteht.

Abb. 1: Künstlich geschaffene Symbole



Abb.2: Erlernte Symbole



Unter System hingegen wird ganz allgemein betrachtet ein „Prinzip, nach dem etwas gegliedert bzw. geordnet wird“ (Brockhaus 2002, Band 3 S.342) verstanden.

Harnad (1990) definiert nun Symbole bzw. Symbolsysteme nach Newell, Pylyshyn, Neumann, Turing, Goedel & Church wie folgt: Unter einem Symbolsystem versteht man eine Reihe an willkürlichen „physikalischen Zeichen“, wie beispielsweise Kratzer auf Papier oder Löcher in einem Band. Diese wurden aufgrund „expliziter Regeln“ – physikalische Zeichen bzw. Zeichenketten - verändert. Der Grundsatz in Bezug auf die Handhabung von Symbolen basiert allerdings lediglich auf der Form und nicht der Bedeutung des Symbolzeichens. Daraus lässt sich ableiten, dass diese Richtlinien lediglich syntaktisch sind und aus Kombinationen und Rekombinationen von Symbolzeichen bestehen. Es gibt primitive atomische Symbolzeichen sowie zusammengesetzte Symbolzeichen (Zeichenketten). Das Gesamtsystem und all seine Einzelteile – die atomischen Zeichen, die zusammengesetzten Zeichen, die syntaktischen Handhabungen sowie die Richtlinien – sind alle semantisch interpretierbar: Dem Syntax kann auf systematische Art und Weise eine Bedeutung zugeschrieben werden, beispielsweise als Abkürzung für Objekte oder der Beschreibung von Sachverhalten. (vgl. Harnad 1990, S.336)

4. Konnektionistische Systeme:

Der Konnektionismus stellt eine Verbindungslehre dar, die sich einerseits mit der Verknüpfung neuronaler Elemente aber auch der Art und Weise inwiefern dadurch höhere Kognitionen erklärt und dargestellt werden können, beschäftigt. Des Weiteren dienen konnektionistische Modelle dazu, Informationen mittels nervenzellartiger Elemente zu verarbeiten, Aktivationen zu sammeln sowie erregende und hemmende Einflüsse auf Andere auszuüben. Somit wird die Information als Aktivationsmuster der neuronalen Elemente dargestellt, wodurch sich die höhere Kognition erklären lässt und sichtbar gemacht wird. (vgl. Anderson 1996, S.146)

Der Konnektionismus wird in der Theorie von Harnad allerdings nur als kognitive Theorie verstanden. Harnad betont, dass „a cognitive theory must stand on its own merits“ (Harnad 1990, S.3); diese eigenen Werte („merits“) sind jedoch davon abhängig, wie gut unser beobachtbares Verhalten erklärt wird. Als solch kognitive Theorie hat der Konnektionismus in der letzten Zeit den symbolischen Ansatz dazu aufgefordert seinen Geist „zu modellieren“. Gemäß dem Konnektionismus stellt nicht die Erkenntnis, sondern die Symbol-Manipulation das dynamische Aktivitätsmuster in einem „mehrschichtigen“ Netzwerk dar. Diese Muster können sich aufgrund interner Netzwerke verändern, woraus ein System resultiert, das dazu in der Lage ist Muster zu erkennen, Probleme zu lösen aber auch motorische Fähigkeiten aufzuzeigen.

Harnad legt allerdings nahe, dass der Konnektionismus, wie er zu Beginn beschrieben worden ist, nicht als symbolisch bezeichnet werden kann. Er begründet dies damit, dass konnektionistische Netzwerke nicht den in Kapitel 3 genannten Kriterien für ein Symbolsystem entsprechen. Aufgrund dessen, dass die Zerlegung bzw. Kombination nach expliziten Regeln der Syntax nicht gegeben ist, ist es nicht möglich konnektionistische Strukturen semantisch zu interpretieren. Ganz einfach verdeutlicht werden kann dies anhand des mathematischen Denkens, da mathematische Formeln oder auch Texte sowohl symbolische wie auch systematische Bedeutung und Inhalte zugesprochen wird; deshalb sind konnektionistische Modelle nicht wirklich dazu geeignet kognitive Leistungen zu erklären. (vgl. Harnad 1990, S.3)

5. The Symbol Grounding Problem

Steven Harnad gilt als Begründer des Symbol Grounding Problems, welches sich auf „die Frage nach der Bedeutung einer „Wissenssprache“ bezieht“ (vgl. Riegler 2000, S.4). Die zentralen Fragen des Symbol Grounding Problems lauten:

„How can the semantic interpretation of a formal symbol system be made intrinsic to the system, rather than just parasitic on the meanings in our heads? How is symbol meaning to be grounded in something other than just more meaningless symbols?“ (Harnad 1990, S. 335).

In anderen Worten ausgedrückt werden Fragen gestellt, die sich darauf beziehen inwiefern die Bedeutung von Symbolen, die definitionsgemäß bedeutungsfrei und einzig auf Grundlage ihrer formalen Eigenschaften verarbeitet sind, in bedeutungsfreie Symbole verankert werden können. (vgl. Riegler 2000, S. 4) Andererseits stellt er sich auch die Frage, wie die semantische Interpretation von einem Symbolsystem intrinsisch vom System erzeugt werden kann, ohne dass dem Symbol die Bedeutung durch den Benutzer zugeschrieben wird.

Des Weiteren stellt das Symbol Grounding Problem die Antwort auf das chinesische Zimmer, einem Gedankenexperiment vom Philosophen John Searle dar, welches wiederum eine Antwort auf den sogenannten Turing-Test liefert. In weiterer Folge soll das Symbol Grounding Problem anhand der eben genannten Verfahren, dem chinesischen Zimmer sowie dem Turing-Test, dargestellt werden.

5.1. Turing-Test

Alan Turing lieferte im Jahre 1950 mit seiner Arbeit „Computing Machinery and Intelligence“ eine rege Erkenntnis in Sachen Computertechnologie. Dabei entwickelte er ein Testverfahren, den Turing-Test, und ging dabei der Frage nach, ob Maschinen in der Lage sind zu denken. (vgl. Lüdi & Lüscher 2007)

Turings anfängliche Intention lag jedoch darin der Frage nachzugehen, ob ein Mensch dazu in der Lage ist durch Kommunikation seinem/ihrem Gesprächspartner das richtige Geschlecht zuzuordnen. Es dauerte nicht lange ehe Turin erkannte, dass dieser Test auch dazu verwendet werden kann die „Qualität der Intelligenz einer Maschine zu analysieren“ (Künstliche Intelligenz – Das Ende der Menschheit). Führt man eine derartige Veränderung der

Testumgebung durch, wodurch die Testperson nicht nur mit einer, sondern mit zwei Personen interagiert – hierbei handelt es sich allerdings bei einer der zu interagierenden Personen lediglich um einen Computer – erhält man den bis dato bekannten Turing-Test. Dieser Test unterscheidet sich im Aufbau nur geringfügig von dem ihm zugrundeliegenden Testverfahren. Die Zielperson erhält die Instruktion mittels virtuellen Gesprächen, die allerdings zeitlich eingeschränkt sind, der Frage nachzugehen, bei welchem der beiden „Personen“ es sich um die reale Person bzw. den Computer handelt. (Künstliche Intelligenz – Das Ende der Menschheit) Dabei geht der Turing-Test von der Annahme aus, dass ein Computer als intelligent bezeichnet werden kann, wenn man die Antwort eines solchen nicht von der einer realen Person unterscheiden kann. (vgl. Lüdi & Lüscher 2007).

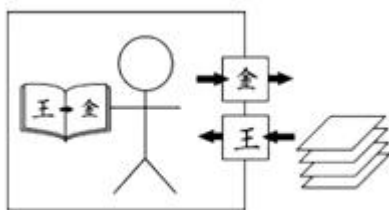
Zum damaligen Zeitpunkt vermutete Turing, dass bis zum Jahre 2000 ein Programm existieren werde, „welches in 30% aller Tests einer Person für 5 Minuten eine menschliche Intelligenz simuliert.“(Künstliche Intelligenz – Das Ende der Menschheit) Leider wurde bis dato noch kein solches Programm entwickelt, wodurch Turing mit seiner Aussage falsch lag.

5.2. Das chinesische Zimmer

John Searle, ein englischer Philosoph führte 1980 als Reaktion auf den Turing-Test von Alan Turing ein Gedankenexperiment durch, welches später unter der Bezeichnung „das chinesische Zimmer“ bekannt wurde. Searle übt Kritik am Turing-Test und behauptet anhand dieses Gedankenexperimentes, dass ein Programm, welches intelligent erscheint, nicht unbedingt intelligent sein muss. (vgl. Kwakwu Ananse)

In erster Linie versucht Searle darzustellen, dass mithilfe eines Computerprogrammes die menschliche Intelligenz in ihren Grundzügen nicht simuliert werden kann.

Das chinesische Zimmer kann wie folgt angesehen werden:



Eine Person, die des Chinesischen nicht mächtig ist – d.h. er/sie kann weder Chinesisch schreiben, noch lesen und auch nicht die Schriftzeichen als solche erkennen -, befindet sich in einem Raum.

Abb.4.

Dieser Person werden durch einen Schlitz in der Wand Geschichten auf Chinesisch zugeschickt. Zudem erhält er/sie Fragen, die sich auf die Geschichten beziehen, welche wiederum auf Chinesisch sind. Als Hilfsmittel ist es der Person erlaubt ein „chinesisches Handbuch“ zu verwenden, wodurch es anhand der eingegangenen Symbole – Geschichte und Fragen - ermöglicht wird eine Antwort auf Chinesisch zu schreiben. Allerdings folgt die Person bei dieser Instruktion ausschließlich den Anweisungen der Anleitung, und versteht somit die Antworten, welche er/sie wiederum durch den Schlitz nach außen schiebt, in keinsten Weise. Der Empfänger dieser Nachrichten ist ein/eine chinesischer/chinesische Muttersprachler/in, der/die nach Empfang der Nachrichten zu dem Schluss kommt, dass sich im Raum drinnen ebenfalls eine Person befinden müsse, die des Chinesischen mächtig ist. (vgl. Searle 1980, S. 428)

Anhand dieser kurzen Instruktion lässt sich auf eindrucksvolle Art und Weise erkennen, dass Searle den Turing-Test quasi „auf den Kopf“ gestellt hat, da der Fokus nicht nur auf das Äußere gelegt wird – im Sinne des Turing Tests würde die Behauptung der Außenstehenden Person, dass die Person im Raum drinnen chinesisch kann, nicht überprüft, sondern als wahr angenommen werden – sondern auch auf das Innere – in diesem Falle, die Person, die sich wie bereits mehrfach gesagt im Inneren befindet und des Chinesischen nicht mächtig ist – gelegt werden. (vgl. Searle 1980, S.447-449)

Seit der erstmaligen Durchführung dieses Gedankenexperimentes sind über 30 Jahre vergangen und beinahe alle Aspekte dieses Modells wurden im Laufe der Zeit von diversen Wissenschaftler, u.a. Harnad, auf welchen in Folge eingegangen wird, widerlegt.

5.3. Chinese Dictionary-Go-Round

Harnad deutet in seinem Artikel darauf hin, dass sich das Symbol Grounding Problem in Bezug auf Searles Chinesisches Zimmer in der Intentionalität äußert. Auch betont Harnad, dass Symbole – Input und Output des Chinesischen Zimmers – sowie die Symbol Manipulationen lediglich auf ihrer Form und nicht wie fälschlicherweise angenommen werden kann auf ihrer Bedeutung basieren. Des Weiteren werden diese systematisch interpretiert und die Frage, was unter einem Symbol System verstanden werden kann, leitet sich einzig von unserer Definition ab. Harnad betont allerdings, dass es keine einheitliche

Interpretation in Bezug auf das Symbolsystem gibt, sondern dass es parasitär bezüglich der Tatsache ist, dass diese Symbole für uns eine Bedeutung haben. Um dies zu verdeutlichen schlage man einfach diverse Bücher auf und suche da nach der Bedeutung von einem gewissen Symbol. Schnell wird man erkennen, dass keine einheitliche Bedeutung dafür existiert, sondern diese sich aus unseren bereits gemachten Vorstellungen ergeben. (vgl. Harnad 1990, S. 338)

Von Harnads eigenem Beispiel des Symbol Grounding Problems existieren zwei Versionen, eine schwierige und eine von der er selbst der Ansicht ist, dass sie unmöglich ist. In der schwierigen Version soll der Proband einzig mit zu Hilfenamen eines chinesischen Wörterbuchs diese Sprache erlernen. Die Reise durch das Wörterbuch würde in etwa einem merry-go-round entsprechen, vorbei an einem bedeutungslosen Symbol zum nächsten, wobei man nie zum Stillstand, im Sinne der Bedeutungsfindung für das Symbol, gelangt. Dieser „trip“ erscheint zwar ausweglos, aber stellt gewiss kein unmögliches Unterfangen dar – man erinnere sich, Harnad selbst bezeichnete diese erste Version als schwierig. Aufgrund dessen, dass jede Person diverse Kenntnisse einer Sprache – Satzbau, Grammatik, etc. – besitzt, wird dieses Vorhaben erheblich erleichtert. Die Unmöglichkeit würde einzig darin bestehen Chinesisch mit Hilfe eines Wörterbuches als erste Sprache zu erlernen. Unter Berücksichtigung dieser Aspekte stellt sich Harnad die Frage, inwiefern es einer Person möglich ist aus einem solchen „symbol/symbol merry-go-round“ (Harnad 1990, S.338) auszusteigen aber auch ob die Bedeutung von Symbolen in etwas anderem als bedeutungslosen Symbolen liegt bzw. liegen kann. Dadurch, dass die Person nur über einen symbolischen Syntax verfügt, aus welchem sich über direktem Weg keine symbolische Bedeutung erkennen lässt, postuliert Harnad in diesem Verfahren das eigentliche Problem der Symbol Grounding zu finden.

Die Standardantwort der Symbolisten, wie beispielsweise von Fodor wäre, dass Symbole ihre Bedeutung erst dann erlangen, wenn das Symbolsystem „richtig“ mit der Welt verbunden worden ist. Harnad erwidert darauf, dass es offensichtlich ist, dass das Problem, die Welt auf die „richtige Art und Weise“ zu verbinden nahezu deckungsgleich mit dem Kognitionsproblem ist. Selbst wenn jede Definition mit der Umwelt auf die „richtige Art“ in einem chinesischen Wörterbuch bzw. im Chinesischen selbst zu finden wäre, würde man trotzdem eine Definition benötigen. Viele Symbolisten glauben, dass Kognitionen autonom funktionierende Module sind, die einzig von peripheren Anordnungen angekoppelt werden müssen, damit sie die Welt mit all ihren Objekten inklusive der Symbole, auf die sie sich

beziehen, sehen können. Bedauerlicherweise, wird die Fähigkeit Objekte aber auch Ereignisse den ihnen zugehörigen Symbolen zuzusprechen unterschätzt oder anders ausgedrückt: das Symbol Grounding Problem wird verharmlost.

6. Die menschlichen Fähigkeiten

Harnad schlägt deswegen vor, man solle zur Bewältigung dieses Problems von einem nonsymbolic/symbolic System ausgehen, das auf zwei elementaren Symbolen, deren Basis zwei Arten von nicht symbolischen Repräsentationen bilden, gründet. Beim Versuch dieses System zu verstehen sollte man sich in erster Linie darauf berufen, wozu der Mensch überhaupt in der Lage ist: er kann unterscheiden, manipulieren, identifizieren, Objekte beschreiben, Beschreibungen erstellen sowie eine Antwort auf diese geben. Dies führt allerdings auch die Frage mit sich, was überhaupt die Voraussetzungen für die eben genannten Fähigkeiten sind?

Damit ein Mensch überhaupt dazu in der Lage ist ein Urteil abzugeben, muss er primär eine Unterscheidung treffen, die sich darauf bezieht, ob zwei gegebenen Inputs ident bzw. kontrovers sind. In diesem Fall wird die Kapazität als Grundlage gesehen, um überhaupt eine Entscheidung zu treffen, was eben dazugehört und was nicht. Um überhaupt dazu in der Lage zu sein, Identifikationen vorzunehmen, muss in erster Linie eine Reihe von äquivalenten Inputs gebildet werden; was wiederum alles auf der Kapazität beruht.

Denken Sie nun an ein Pferd. Der Mensch ist dazu in der Lage verschiedene Pferdearten auseinanderzuhalten aber auch die Gemeinsamkeiten zwischen diesen zu erkennen, bezeichnet man dies als Unterscheidung. Wenn man nun ein Pferd sieht und dieses als solches erkennt – und nicht als Maultier, Esel oder dergleichen -, handelt es sich allerdings um Identifikation. (vgl. Harnad 1990, S.338 – 339)

In Bezug darauf stellt sich Harnad die Frage, welche Art innerer Repräsentation benötigt wird, damit wir überhaupt in der Lage sind zwischen diesen zwei Fähigkeiten zu unterscheiden.

6.1 Ikonische und kategoriale Repräsentationen

Ikonische Repräsentationen bilden die Grundlage dafür, dass diverse Sachverhalte voneinander unterschieden bzw. Gemeinsamkeiten erkannt werden können. (vgl. Harnad 1990, S. 339) Sie können auch als „geistige Bilder“, welche von uns in diversen Situationen gemacht werden, verstanden werden. Zu beachten ist lediglich, dass diese Bilder, die auch in anderen Wahrnehmungssinnen verankert sein können, eine „nachvollziehbare Ähnlichkeit mit dem von ihnen bezeichnenden Objekt haben“ und keineswegs nur visuell sind. (vgl. Jank & Meyer 2002)

Somit werden Symbole bzw. Vorstellungen von einem Pferd benötigt um dieses von anderen Pferden zu unterscheiden. Es existieren in unserer Welt jedoch nicht nur eindeutig zuordenbare Repräsentationen, weshalb eine Identifikation, bei welcher der Sachverhalt bekannt sein muss, nicht möglich ist. (vgl. Harnad 1990, S.339)

Damit Sachverhalte erst identifiziert und somit im weiteren Sinne benannt werden können, müssen unsere Vorstellungen des jeweiligen Sachverhalts auf wenige Merkmale reduziert und anhand derer in Kategorien geordnet werden. Dabei sind viele der uns bekannten Kategorien angeboren und weitere können im Verlauf des Lebens durch Erfahrung angeeignet werden; ein Beispiel dazu stellt die Kategorie „Pferd“ dar.

Es soll allerdings nicht außer Acht gelassen werden, dass die ikonische sowie die kategoriale Repräsentation von einem Sachverhalt lediglich sensorisch – sie werden in Form von „Fotos“ in unseren Köpfen gespeichert - und keineswegs symbolisch sind. Die Frage, inwiefern die Vorstellungen mit der realen Welt verbunden sind, wird demnach außer Acht gelassen ebenso wie die der semantischen Interpretierbarkeit und deren Richtigkeit. (vgl. Harnad 1990, S.340)

6.2 Symbolische Repräsentationen

Mittels symbolischer Repräsentation können Sachverhalte anhand von Kategorien, Regeln und Begriffe, unabhängig vom konkreten Gegenstand bzw. dem eigenen Tun, abstrakt erfasst werden: Die Zahl 5 hat nichts mit fünf beliebig gewählten Gegenständen zu tun, ebenso wie der Notenkopf im Violinschlüssel nichts mit dem auf einem Instrument gespielten x-beliebigen Ton zu tun hat. (vgl. Jank & Meyer 2002)

Auch kategoriale Repräsentationen können das Dahinterstehende nicht interpretieren. Zwar können sie eine Auswahl an Objekten treffen, die sie benennen, aber die Bezeichnungen besitzen nicht die Symboleigenschaften sowie die bereits beschriebenen Kriterien für Symbole und Symbolsysteme. Man erinnere sich: damit etwas als systematisch bezeichnet werden kann, muss es kombiniert und wieder verändert werden können. „Pferd“ stellt nur eine zuverlässige Bezeichnung für eine bestimmte Kategorie dar, allerdings gibt es nichts, das uns sagt, dass es sich hierbei auch wirklich um ein „Pferd“ handelt – die Bezeichnung wurde in der Gegenwart gehandhabt und ein jeder/ein jede weiß, was damit gemeint ist und wenn wir das Wort „Pferd“ hören assoziieren wir es mit diversen Dingen. Was wäre nun erforderlich, um die anderen systematischen Eigenschaften zu erzeugen? Es wäre lediglich von Nöten, dass die Bezeichnung in der Systematik verankert wird, zusammen mit anderen Aussagen, die sich auf eben diese Kategorie beziehen. Als Beispiel dazu

1. Denken Sie nun an die Bezeichnung „Pferd“. Diese basiert auf ikonischen sowie kategorialen Repräsentationen und im Laufe unseres Lebens haben wir gelernt, Pferde auf Grundlage ihrer sensorischen Projektion zuverlässig zu unterscheiden und zu identifizieren.
2. Bei der Bezeichnung „Streifen“ lassen sich ähnliche Schlüsse ziehen.
3. Zebra = Pferd & Streifen

Die mentale Repräsentation eines Zebras ergibt aus der Kombination der Symbole „Pferd“ und „Streifen“ eine Zeichenkette. Da diese beiden Symbole durch sowohl bildliche als auch kategoriale Repräsentationen in unseren Köpfen verankert sind, erhält das Zebra diese „Verankerungen“, wodurch es laut Harnad zu einer „grounded symbolic representation“ (Harnad 1990, S.340) wird. So könnte praktisch ein jeder/ jede – auch wenn er/sie zuvor noch nie ein Zebra gesehen haben sollte – durch die verankerten Repräsentationen der Symbole „Pferd“ und „Streifen“ mittels der symbolischen Repräsentation diese nicht nur identifizieren sondern auch benennen. Somit stellt die Fähigkeit, Sachverhalte der Welt mithilfe von ikonischen und kategorialen Repräsentation zu unterscheiden und zu benennen auch die Fähigkeit dar, diese anhand von symbolischen Repräsentationen zu treffen.

6.3 Die komplementäre Rolle des Konnektionismus

Beruft man sich jedoch auf die Entstehung diverser Repräsentationen in unseren Köpfen sowie die Auswahl an zentralen Merkmalen zur Kategorienbildung, bezieht man sich auf den Konnektionismus: Unsere Vorstellungen sind das Produkt konnektionistischer Netzwerke, die eine Beziehung zwischen den verschiedenen Sachverhalten und den dahinterstehenden Dingen herstellen. Auch unterstützen sie uns bei der Findung der für eine Kategorienbildung wesentlichen Merkmale. (vgl. Harnad 1990, S.340)

Diese Rolle als Fundament der Begriffe und somit Lösung des Symbol Grounding Problems im Symbolischen Ansatz des Denkens beendet deren alten Wettstreit und ergibt ein Modell, das Harnad als „Hybrides System“ bezeichnet. Dieses ist dem rein konnektionistischen überlegen, da es die Fähigkeit des Menschen beinhaltet, mit semantisch interpretierbaren Symbolen zu agieren, sowie dem rein symbolischen Ansatz, dem die sichere Verankerung der richtigen Bedeutungen der Symbole fehlt.

7. Das Hybride System

Das beschriebene System funktioniert natürlich nur, wenn die Zuschreibung der Bedeutungen „bottom up“ erfolgt, das heißt dass zunächst die elementaren Begriffe durch die sensorischen Repräsentationen ihre Bedeutungen erhalten und daraufhin die höheren Symbole diese übernehmen. Hier zeigt sich auch der fundamentale Unterschied zum Computer, dessen Software seine Bedeutungen nicht auf diesem Wege erhält, sondern immer nur auf der symbolischen Ebene operiert.

Die Symbole eines eben beschriebenen mit einer Umwelt interagierenden Systems können allerdings nicht willkürlich geformt sein, wie die eines rein syntaktischen Systems, da sie von ikonischen und kategorialen Repräsentationen der Realität abhängig sind. Auch wenn dies noch nicht näher untersucht wurde, so müsste es laut Harnad im Verhalten zu belegen sein: Sowohl die Kriterien eines Symbolsystems müssten vorhanden sein – vgl. Kapitel 3 -, als auch die Fähigkeiten zu unterscheiden, zu identifizieren und Symbole den richtigen Objekten oder Zuständen zuzuschreiben. Kämen diese Untersuchungen zu einem positiven Ergebnis, würde das natürlich noch nicht beweisen, dass das Denken tatsächlich so funktioniert, aber es wäre das bis dato beste existierende Modell. (vgl. Harnad1990, S.341)

8. Kann künstliche Intelligenz das Symbol Grounding Problem überwinden?

Um auf die Frage zurückzukommen, ob eine künstliche Intelligenz nun wirklich die Symbole verstehen kann mit der es operiert, so kann dies nach Harnad eben nur dann der Fall sein, wenn diese mit ihrer Umwelt interagieren kann. Tatsächlich existieren bereits zahlreiche Versuche, Roboter zu bauen, die von ihrer Umwelt lernen und kommunizieren können. Diese werden auch als autonome Agenten bezeichnet.

Luc Steels schreibt in seinem Forschungsbericht „The Symbol Grounding Problem has been solved. So what`s next?“ aus dem Jahr 2006, dass bereits solche Roboter existieren. Zunächst erwähnt er Nilsson`s „Shakey“ aus den Siebziger, dem man befehlen konnte, einen gewissen Ort im Nachbarraum aufzusuchen. Er verstand demnach Symbole (den Ort) als Teil einer realen Welt; allerdings waren diese von Menschen einprogrammiert, und demzufolge sind sie nicht auf die Art und Weise entstanden, wie Harnad dies gefordert hat. Mittlerweile gibt es jedoch schon Roboter, die auch dies können, und somit ihre Symbolsysteme von der Umwelt lernen und auf ihre Richtigkeit überprüfen können. Ist das Symbol Grounding Problem damit gelöst? Hier erwähnt Steels, dass ein weiteres Problem darin besteht, dass auch für uns Menschen manchmal Symbole unterschiedliches bedeuten können, die Bedeutungen dabei aber alle richtig und somit individuell sind; demzufolge müssten eigentlich alle Symbole permanent überdacht werden. Menschen verstehen sich aber trotzdem, da sie über eine Fähigkeit verfügen, die Steels „alignment“ nennt und die aus unserer Fähigkeit zur theory of mind entsteht.

Im Allgemeinen könnte es sein, dass gewisse Teile der Intelligenz eine gewisse Hardware erfordern, die in der Art und Weise bisher allerdings nur im menschlichen Gehirn existiert, wobei es hierfür noch keinerlei Beweise gibt. Möglicherweise limitiert auch die Überlegenheit des menschlichen Körpers die Fähigkeiten einer lediglich in einem Roboter installierten künstlichen Intelligenz.

Trotz dieser immensen Fortschritte steckt die Forschung hier noch in den Kinderschuhen: Technik und Theorie werden sich mit Sicherheit in den nächsten Jahren so stark verbessern, dass neue Erkenntnisse nur eine Frage der Zeit darstellen. Spannend wird sicherlich die Auswirkung der möglichen Erkenntnis, dass wir Menschen den Maschinen vielleicht doch nicht überlegen sind, wie bereits schon viel früher in den Filmen Matrix und Terminator

prophezeit wurde. Eventuell tut sich auch von religiöser Seite Widerstand gegen eine Forschung auf, die irgendwann vielleicht menschliches Denken - die Basis der „Seele“ - komplett auf biologische Mechanismen zurückführen kann.

Literaturverzeichnis:

Anderson, J. R. (1996). *Kognitive Psychologie*. 2. Aufl. Heidelberg: Spektrum Akademischer Verlag

Baumann, R., Gumbel, G., Jacob, R., Kratzmaier, P., Schneider-Nicolay, B., Zimmer, E. (2002): *Der Brockhaus von A-Z*. In 3 Bänden. Sonderausgabe. Augsburg: Weltbild.

Beckermann, Ansgar (2008): *Analytische Einführung in die Philosophie des Geistes*. 3. Aufl. Berlin: Walter de Gruyter GmbH

Harnad, S. (1990) *The Symbol Grounding Problem*. Physica D 42: 335-346.

Internet 1: *Psychologie Lexikon*. Verfügbar unter:

<http://www.psychology48.com/deu/d/symbol/symbol.htm> (Stand 2011-04-14)

Internet 2: *Psychologie Lexikon*. Verfügbar unter:

<http://www.psychology48.com/deu/d/kognitive-wende/kognitive-wende.htm> (Stand 2011-04-14)

Jank, Werner & Meyer, Hilbert (2002): *Didaktische Modelle*. Verfügbar unter:

<http://beat.doebe.li/bibliothek/w01893.html>

Künstliche Intelligenz – Das Ende der Menschheit. Verfügbar unter: http://www.uni-magdeburg.de/iew/web/studentische_projekte/ws04/berger/turing-test.htm (Stand 2011-04-15)

Kwaku Ananse. Verfügbar unter: <http://kwakuananse.twoday.net/stories/5934718/> (Stand 2011-04-14)

Lercher, Silvia & Heinisch, Alexandra (2009): *The Symbol Grounding Problem*.

Semesterarbeit im Rahmen des Forschungsseminars: Neuere psychologische Fachliteratur.

Verfügbar unter:

http://www.uibk.ac.at/psychologie/mitarbeiter/leidlmair/the_symbol_grounding_problem_seminararbeit.pdf (Stand 2011-04-14)

Luc Steels (2006), *The Symbol Grounding Problem has been solved. So what's next?*

Verfügbar unter: <http://ecagents.istc.cnr.it/dllink.php?id=267&type=Document> (Stand 2011-04-28)

Lüdi, Gregor & Lüscher, Martin (2007): *Können Maschinen denken?* Verfügbar unter:

<http://www.natur-struktur.ch/ai/turingtest.html> (Stand 2011-04-15)

Lexikon für Psychologie und Pädagogik. Verfügbar unter:

<http://lexikon.stangl.eu/182/behaviorismus/> (Stand 2011-04-14)

Rickheit, Gert, Herrmann, Theo & Deutsch, Werner (2003): *Psycholinguistik: Ein internationales Handbuch*. Berlin/New York: Walter de Gruyter

Riegler, Alexander (2000): *Können wir das Problem der Echtzeit lösen?* In: Edlinger, K. & Fleck, G. (eds). *Artificial Life: Organismisch-systemische Perspektiven künstlichen Lebens*. (Arbeitstitel) München: Peter Lang

Searle, J. R. (1980). *Minds, brains, and programs*. Behavioral and Brain Sciences 3 (3): 417-457.

Schulke, Jutta (2010): *Cognitive Science – Bachelor of Science*. Verfügbar unter:

http://www.uni-osnabrueck.de/160_12030.html (Stand 2011-04-14)

Welt der Wunder. MSN Wissen, Verfügbar unter: <http://weltderwunder.de.msn.com/mensch-und-natur-gallery.aspx?cp-documentid=149940623&page=7> (Stand 2011-04-23)

Werner Jank, Hilbert Meyer (2002), Didaktische Modelle im Text Lernen und Entwicklung. Verfügbar unter: <http://beat.doebe.li/bibliothek/w01893.html> (Stand 2011-05-01)

Wörterbuch und Enzyklopädien auf der Akademik. Verfügbar unter:

http://de.academic.ru/dic.nsf/dewiki/1273000#cite_note-0 (Stand 2011-04-14)

Abbildungsverzeichnis:

Bild des Titelblattes: CCSG. Concordia Cognitive Science Group. Verfügbar unter:

<http://linguistics.concordia.ca/ccsg/> (Stand 2011-04-15)

Abb.1: hot plot. Aufkleber und mehr. Verfügbar unter:

http://www.hotplot.de/shop_new/index.php?cat=c45_Rauchen-Verboten.html (Stand 2011-04-15)

Abb.2: Greiner, Carmen (unbekannt): Psychologie. Verfügbar unter: [http://greiner-](http://greiner-brg.members.cablelink.at/index_psycho.html)

[brg.members.cablelink.at/index_psycho.html](http://greiner-brg.members.cablelink.at/index_psycho.html) (Stand 2011-04-15)

Abb.3: Wörterbuch und Enzyklopädien auf der Akademik. Verfügbar unter:

http://de.academic.ru/dic.nsf/dewiki/1273000#cite_note-0 (Stand 2011-04-14)

Abb.4: IuG – Künstliche Intelligenzen. Das chinesische Zimmer. Verfügbar unter:

http://www.informatik.uni-oldenburg.de/~iug08/ki/Grundlagen_Chinese_Room.html (Stand 2011-04-15)