

The Symbol Grounding Problem



Semesterarbeit im Rahmen des Forschungsseminars:

Neuere psychologische Fachliteratur

Dozent: Dr. Karl Leidlmair

VO-Nr.: 720 584

SS 2009

Studienkennzahl: C298

Vorgelegt von:

Silvia Lercher

0515849

Alexandra Heinisch

0616701

Inhaltsverzeichnis

1 Einleitung	3
2 Vom Behaviorismus zum Kognitivismus	4
3 Symbolsysteme	6
4 The Symbol Grounding Problem.....	8
4.1 Stevan Harnad.....	9
4.2 Turing Test	9
4.3 Chinesisches Zimmer	10
4.4 Merry-Go-Round.....	11
4.5 Repräsentation	13
4.5.1 Ikonische Repräsentation.....	13
4.5.2 Kategoriale Repräsentation.....	13
4.5.3 Symbolische Repräsentationen.....	14
5. Lösungsansätze	15
5.1 From the Ground up	15
5.2 Autonome Agenten.....	16
6. Schlussworte	19
Literaturverzeichnis	20
Abbildungsverzeichnis.....	21

1 Einleitung

Im Rahmen des Forschungsseminars Neuere psychologische Fachliteratur haben wir von Dr. Karl Leidlmair die Aufgabe bekommen das Thema „The Symbol Grounding Problem“ näher zu bearbeiten.

The Symbol Grounding Problem wurde von Steven Harnad definiert, welches die Antwort auf das Gedankenexperiment des Chinesischen Zimmers von John Searle darstellt. Zentrale Fragen des Symbol Grounding Problems (zitiert nach Harnad, 1990) 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?“ (S. 335).

Mit anderen Worten ausgedrückt stellt er sich einerseits die Frage „Wie die Bedeutung von Symbolen für Systeme intrinsisch vom System selbst erzeugt werden kann, anstatt durch die Zuschreibung der Bedeutung durch den Benutzer?“ und andererseits stellt er sich auch die Frage „Wie die Bedeutung von Symbolen in etwas anderem als in bedeutungslosen Symbolen liegen kann?“.

Diesen zentralen Fragen werden wir in der nachfolgenden Arbeit nachgehen und geben zusätzlich einen Überblick über die geschichtlichen Zusammenhänge, erläutern Symbolsysteme und gehen weiters noch auf das Chinesische Zimmer und den Turing Test ein. Enden wird unsere Arbeit mit einigen Vorschlägen zur Lösung des Symbol Grounding Problems.

2 Vom Behaviorismus zum Kognitivismus

Aus der Geschichte.....

Das Symbol Grounding Problem ist keine Problematik, die seit je her diskutiert wurde (Vgl. Wikipedia). Bis in die 1950er Jahre, als der Behaviorismus noch die vorherrschende Denkströmung war, wurden psychische Phänomene und mentale Prozesse aus der psychologischen Forschung völlig ausgeklammert. Jegliches Verhalten des Menschen (und der Tiere) konnte als ein einfaches Reiz-Reaktionsmuster beschrieben werden, der Mensch selber wurde als eine Art „Black-Box“ verstanden. So war rein das beobachtbare Verhalten des Menschen Gegenstand der Forschung.

Es ist nicht weiter verwunderlich, dass die Tatsache, dass innerpsychische Vorgänge keine Beachtung fanden, bald auf heftige Kritik stieß.

In den Jahren 1960/1970 fanden erstmals auch unbeobachtbare Aspekte des menschlichen Verhaltens, wie mentale Zustände und Überzeugungen, Beachtung. Vor allem Avram Noam Chomsky (Vgl. Wikipedia), ein Professor für Linguistik, leistete wichtige Beiträge und Kritik am Behaviorismus.

Der Umbruch vom Behaviorismus hin zu einer kognitiven Denkströmung wird als *Kognitive Wende* verstanden (1956).

Kognitive Fähigkeiten wie Wahrnehmung, Denken, subjektive Interpretationen und Sprache waren nun neben bestimmten Einflüssen der Vererbungslehre Gegenstand der Psychologischen Forschung.

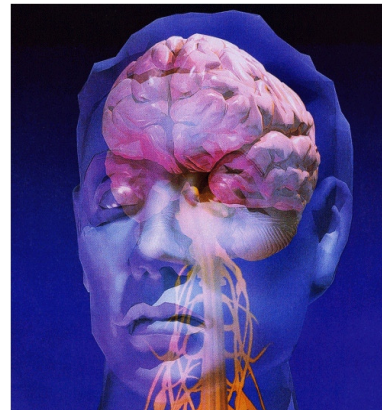


Abbildung 1: Kognitive Wende

Es ging nun darum, psychische Prozesse auf einer formalen bzw. abstrakten Ebene zu beschreiben.

Eine bedeutende Rolle für die Entstehung der kognitiven Wissenschaften spielte natürlich auch die Entwicklung des Digitalcomputers zu dieser Zeit. Der Digitalcomputer wurde von Kognitionswissenschaftlern als Modell für das menschliche Gehirn verstanden, das sogenannte „Computermodell des Geistes“, wo das menschliche Gehirn als informationsverarbeitendes System beschrieben wurde.

Ebenso wie eine formal/abstrakte Beschreibung der Computer-Software möglich ist, ging man nun davon aus, dass man den menschlichen Geist ebenfalls auf abstrakter Ebene beschreiben kann, ohne direkt beobachtbare Verhaltensweisen zu berücksichtigen.

Eine wichtige Rolle innerhalb der Kognitionswissenschaft spielt die „Repräsentationale Theorie des Geistes“ von Jerry Alan Fodor, der die These vertritt, dass das menschliche

Gehirn vergleichbar mit der Hardware eines Computers sei und der menschliche Geist analog zur Software eines Computers beschrieben werden könne, und zwar unabhängig von der Hardware. Kurz: Man kann den menschlichen Geist beschreiben, ohne das Gehirn und seinen neuronalen Aufbau zu kennen (Vgl. Wikipedia).

Fodor (1975) vertritt außerdem die Annahme, dass der menschliche Geist eine Art „Symbolsystem“ ist, und dass dieses „Symbolsystem“ durch Denkprozesse bearbeitet wird. Der menschliche Geist arbeitet mit mentalen Repräsentationen, die durch bestimmte Regeln zu Gedanken werden, ähnlich jeder anderen Sprache, wo bestimmte Wörter zu Sätzen und somit zu Bedeutungen zusammengesetzt werden. Zum Beispiel in der Künstlichen Intelligenz Forschung wurde empirisch nachgewiesen, dass durch Manipulation der Symbolsysteme komplexe Verhaltensweisen entstehen und so erklärt werden können. (Siehe Harnad, 1990).

In späteren Entwicklungen der Kognitionswissenschaften wurde das mechanistische Modell des Geistes – das Computermodell des Geistes – stark kritisiert, v.a. weil es durch die modernen Entwicklungen der Neurokognitionswissenschaften nun möglich war auch das Gehirn (die „Hardware“) genau zu beschreiben und somit eine Beschreibung des Geistes unabhängig vom Gehirn ungenau und die Trennung von „Hardware“ und „Software“ bzw. von Geist und Gehirn unnötig und unbrauchbar schien. So verzichtet zum Beispiel der Konnektionismus auf die Trennung von „Software“ und „Hardware“.

Weitere Strömungen der Kognitionswissenschaften sind unter anderem Artificial life, Artificial intelligence, Kognitionspsychologie, Imbodyment, und Neuronale Netze (Vgl. Wikipedia).

Heute steht die Kognitionswissenschaft in engem Zusammenhang mit den Neurowissenschaften und arbeitet mit neuronalen Netzen. Sprachbeherrschung und die Fähigkeit des Menschen Sprachen zu erlernen, ist im Bereich der Kognitionswissenschaft sehr interessant. Viele Gedanken könnte man ohne Sprache nicht denken, und viele Probleme wären ohne Sprache unlösbar. Vor allem aber ist es interessant, wie Wörter, Sätze, und Sprache insgesamt zu ihrer Bedeutung kommen.

3 Symbolsysteme

Damit wir uns dem Symbol Grounding Problem, wie es von Stevan Harnad beschrieben wurde, nähern können, müssen wir zuerst klären was denn überhaupt ein Symbol ist und was ein Symbolsystem ist.

Symbol kommt aus dem gr. Lat. und bedeutet soviel wie „Kennzeichen oder Zeichen (Vgl. Duden, 2005).

1. In der Antike ein durch Boten überbrachtes Erkennungs- oder Beglaubigungszeichen zwischen Freunden
2. Sinnbild
3. Ausdruck des Unbewussten, Verdrängten in Worten, Handlungen, Traumbildern und ähnlichem
4. Christliches Tauf- oder Glaubensbekenntnis
5. Zeichen, das eine Rechenanweisung gibt
6. Zeichen oder Wort zur Darstellung oder Beschreibung einer Informationseinheit oder Operation
7. Zeichen für eine physikalische Größe



Abbildung 2: Darstellung von Symbolen

Definition aus Wikipedia: Der Terminus Symbol aus dem Griechischen: *Etwas Zusammen-gefügt*, oder auch Sinnbild, wird im Allgemeinen für Bedeutungsträger, Zeichen, Wörter, Gegenstände, Vorgänge und vieles mehr verwendet, die eine Vorstellung meinen, von etwas, das nicht gegenwärtig sein muss.

Harnad (1990) definiert Symbole bzw. Symbolsysteme nach Newell, Pylyshyn, Neumann, Turing, Goedel & Church wie folgt:

Ein Symbolsystem ist eine Reihe zufälliger „Physikalischer Zeichen“ wie Striche auf einem Papier, Löcher in einem Band u.ä. (z.B. folgendes: $\neg\cup\cup\Xi\Gamma\cup\cup\cup\cup\cup\cup\cup$, also eine Zeichenkette bestehend aus mehreren Zeichen/ Symbolen).

Diese Zeichen werden mit Hilfe von expliziten Regeln bearbeitet bzw. manipuliert. Die Regeln sind ebenfalls Zeichenketten.

Diese regelgeleitete Bearbeitung bzw. Manipulation der Zeichenketten beruht lediglich auf der Gestalt bzw. Form der einzelnen Zeichen (Symbole), jedoch nicht auf ihrer Bedeutung. Somit spielt nur die Syntax eine Rolle, also die „Grammatik“, die Form, Ordnung und Beziehung der Zeichen oder Zeichensysteme zueinander.

Die Symbolmanipulation besteht also aus einer regelgeleiteten Kombination und Veränderung der Zeichen / Symbole. Dabei unterscheidet Harnad primitive, einfache Zeichen und Symbole von kombinierten Symbolketten.



Abbildung 3: Beispiel für Zeichen: Hieroglyphen

Das gesamte Symbolsystem mit all seinen einzelnen, primitiven Zeichen aber auch mit all seinen kombinierten und verknüpften Zeichenketten, die syntaktischen Veränderungen bzw. Manipulationen sowie alle Regeln sind „semantisch interpretierbar“. D.h., dass die Syntax systematisch auf Bedeutungen bzw. auf dahinterstehende Inhalte und Sachverhalte der realen Welt übertragen werden können. Auf Objekte, aber auch auf aktuelle Zustände und andere Sachverhalte (Siehe Harnad, 1990).

Nach Fodor's (Siehe 1975) „Symbolic model of mind“ sind genau diese Symbolketten die mentalen, innerpsychischen Phänomene, sprich Kognitionen.

Die expliziten Regeln der Symbolmanipulation sind Teil eines formalen Systems, welches in seine Teile zerlegt werden kann. Die Symbolmanipulation ist rein formal, (Abhängig von der Syntax bzw. der Form der Zeichen) und das System muss als Ganzes semantisch interpretierbar sein, denn ein einzelner Teil kann keine Bedeutung haben. Somit müssen Symbole immer systematisch sein (Vgl. Harnad, 1990).

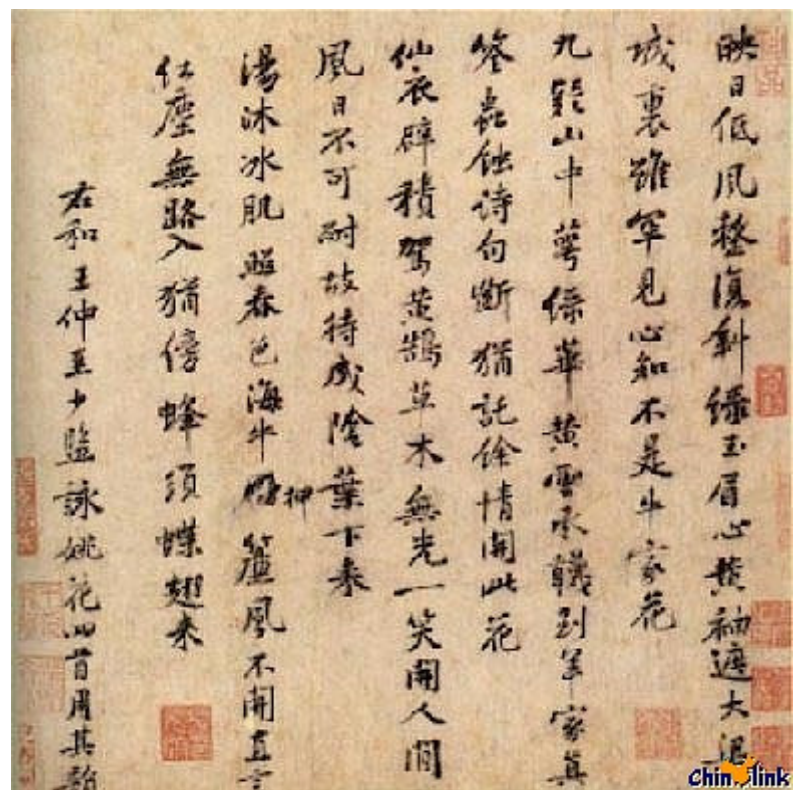


Abbildung 4: Syntax und Semantik

Die Tatsache, dass Verhalten als regelgeleitet verstanden werden kann, muss aber nicht bedeuten, dass es von Symbolregeln beherrscht ist. Semantische Interpretation muss immer mit expliziten Repräsentationen, formalen, syntaktischen Veränderungen sowie einer Systematik verbunden werden, um als symbolische Interpretation verstanden zu werden.

Konnektionistische Systeme

Im Konnektionismus wird ein System bzw. das Verhalten eines Systems z.B. in der künstlichen Intelligenzforschung als eine Kombination vieler einfacher Einheiten, die untereinander vernetzt sind, verstanden. Diese Einheiten „kommunizieren“ miteinander, sind lern- und anpassungsfähig. Ein solches konnektionistisches System ist z.B: das menschliche Gehirn, aber auch künstliche neuronale Netze (Siehe Wikipedia).

Gemäß dem Konnektionismus (Vgl. Harnad, 1990) sind Kognitionen nicht lediglich Symbol-Manipulation, sondern dynamische Strukturen und Aktivitäten in „verknüpften Netzwerken“. So entsteht ein lernfähiges System welches Problemlösefähigkeiten und motorische Fertigkeiten aufzeigt.

Nach Harnad (1990) kann man den oben kurz beschriebenen Konnektionismus allerdings nicht als „symbolisch“ bezeichnen, da die konnektionistischen Netzwerke nicht die obengenannten Kriterien eines Symbolsystems erfüllen. Konnektionistische Strukturen können nicht systematisch semantisch interpretiert werden, da keine Zerlegung oder Kombination nach expliziten Regeln der Syntax gegeben ist. Sehr viele unserer Verhaltensweisen, Kognitionen, sprachliche Kompetenzen und viele weitere Fähigkeiten wie logisches, mathematisches Denken (einfache Sätze, aber auch mathematische Formeln, Operationen, ...) haben aber symbolische und systematische Bedeutungen und Inhalte. Deshalb scheinen konnektionistische Modelle nicht ideal, um solche kognitive Leistungen zu erklären.

4 The Symbol Grounding Problem

Wie wir in der Einleitung schon angeführt haben wurde das „Symbol Grounding Problem“ von Steven Harnad definiert, dessen Kernaussage unter anderem aus der Frage besteht „Wie Wörter ihre Bedeutung erhalten beziehungsweise was Bedeutungen überhaupt sind?“. Das Symbol Grounding Problem stellt genau genommen die Antwort auf das Chinesische Zimmer dar, das von John Searle entwickelt wurde, welches wiederum die Antwort auf den Turing Test darstellt. Aus diesem Grund werden wir sowohl das Chinesische Zimmer als auch den Turing Test näher erläutern und weiters noch Angaben zur Person Steven Harnad geben.

4.1 Stevan Harnad

Stevan Harnad, dessen ursprünglicher Name Hernád István lautet, wurde 1945 in Budapest, Ungarn geboren. Er studierte an der McGill University, sowie an der Princeton University, wo er auch graduierte. Zurzeit hat er Professuren an der University of Southampton, sowie an der Université du Québec in Montreal inne. Zusätzlich wirkt er noch an der Hungarian Academy of Sciences. Seine Forschungsschwerpunkte liegen im Bereich Kategorisierung, Kommunikation und Kognition.



Abbildung 5: Stevan Harnad

Außerdem war Steven Harnad der Begründer und Redakteur von Behavioral and Brain Science, Psycholology and CogPrints (Siehe Wikipedia).

4.2 Turing Test

Der Turing Test geht auf Alan Mathison Turing zurück, der damit der Frage: „Can machines think?“ nachgeht.

Alan Turing beschreibt in seinem Artikel „Computing Machinery and Intelligence“ (1950) den Turing Test, als ein Spiel, als das sogenannte „imitation game“, in dem drei Personen miteinander in Interaktion stehen. Dabei sind eine Frau (A), ein Mann (B) und eine dritte Person (C), unabhängig vom Geschlecht beteiligt.

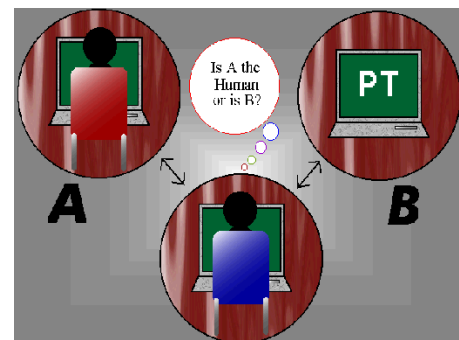


Abbildung 6: Darstellung des Turing Tests

Die dritte Person, die sich getrennt von den anderen Beiden befindet, hat nun die Aufgabe durch Kommunikation herauszufinden welchem Geschlecht die Person A beziehungsweise die Person B angehört. Um diese Aufgabe lösen zu können stellt die dritte Person der Person A und der Person B Fragen, die diese zu beantworten haben. Jedoch erhält der Fragesteller (C) die Antworten nicht mündlich, sondern schriftlich beziehungsweise über einen Computer.

Person A hat dabei die Aufgabe die Person C zu unterstützen und Person B hat die Aufgabe Person C in die Irre zu führen. Ein Auszug eines typischen Gesprächs könnte lauten:

C: „Will B please tell me the length of his or her hair?“

B (Mann): “My hair is shingled and the longest strands are about nine inches long.”

Der eigentlichen Frage, der Turing aber nachgehen möchte lautet: „Was passiert wenn eine Maschine, also ein Computer, die Rolle von A oder B übernimmt?“. Wird der Fragesteller in der Lage sein den Menschen vom Computer zu unterscheiden? Gelingt ihm dies nämlich nicht, hat die Maschine den Turing Test erfolgreich bestanden.

Turing (1950) äußert sich dazu wie folgt:

„I believe that in about fifty years' time it will be possible to programme computers, with a storage capacity of about 10^9 , to make them play the imitation game so well that an average interrogator will not have more than 70 per cent chance of making the right identification after five minutes of questioning.“ (S. 442).

Jedoch hat sich diese optimistische Vorhersage bis jetzt noch nicht erfüllt und bleibt damit ein guter Ansatz für weitere Forschung.

4.3 Chinesisches Zimmer

Das Chinesische Zimmer ist ein Gedankenexperiment von John R. Searle, der damit unter anderem auf den Turing Test von Alan Turing und der „theory of mind“ antwortet. Vorgestellt wurde das Chinesische Zimmer erstmals 1980 in Searle's Artikel „Minds, brains and programs“.

Das Chinesische Zimmer muss man sich wie folgt vorstellen (Vgl. Searle, 1980): Eine Person, die dem Chinesischen in Wort und Schrift nicht mächtig ist wird in einen Raum gebeten, in dem sie Chinesische Schriftzeichen erhält (Input), die in eine bestimmte Reihenfolge gebracht werden müssen und weitergeleitet werden müssen (Output).

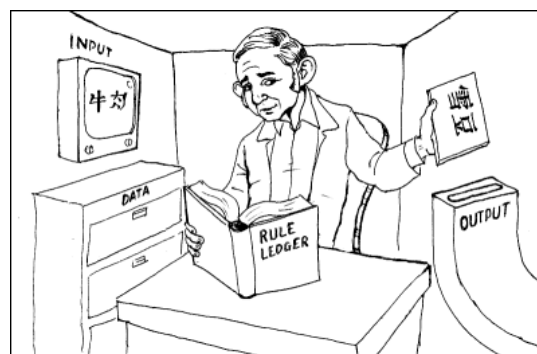


Abbildung 7: Darstellung des Chinesischen Zimmers

Einziges Hilfsmittel dabei stellt ein Buch mit Regeln dar, dass in der jeweiligen Muttersprache verfasst wurde. Die chinesischen Schriftzeichen werden nach dem Prinzip der Schrift, nach dem Prinzip der Frage und dem Prinzip der Geschichte geordnet und an einen weiteren Experimentteilnehmer, dessen Muttersprache Chinesisch ist, weitergeleitet. Aufgrund der Sinnhaftigkeit der Antworten kommt dieser zu dem Schluss dass sein Gesprächspartner, ebenfalls dem Chinesischen mächtig ist, obwohl bei diesem kein Verständnis für Geschichte, Frage oder Antwort vorliegt.

Die alt bekannte Frage: Können Maschinen denken?, die Turing als Ausgangspunkt seiner Forschung diente, beantwortet Searle ganz klar mit einem Ja, sie können.

Jedoch geht er noch weiter und fragt: Kann ein Artefakt, etwas vom Menschen künstlich erzeugtes denken? oder Kann ein digitaler Computer denken?. Beide Male lautet die Antwort Ja, sie können. Doch dann stellt Searle (1980) die eigentliche Frage: *„But could something think, understand, and so on solely in virtue of being a computer with the right sort of program? Could instantiating a program, the right program of course, by itself be a sufficient condition of understanding?“*(S. 428). Und diesmal lautet die Antwort Nein. Aber warum? *“Because the formal symbol manipulations by themselves don't have any intentionality; they are quite meaningless; they aren't even symbol manipulations, since the symbols don't symbolize anything. In the linguistic jargon, they have only a syntax but no semantics. Such intentionality as computers appear to have is solely in the minds of those who program them and those who use them, those who send in the input and those who interpret the output”* (S. 429).

4.4 Merry-Go-Round

Steven Harnad bezieht sich in seinem Artikel „The Symbol Grounding Problem“, der 1990 veröffentlicht wurde, explizit auf das Chinesische Zimmer von John Searle und startet seine Erklärung dazu mit der Aussage, dass sich das symbol grounding problem eigentlich im Gedankenexperiment des Chinesischen Zimmers als das Problem der Intentionalität äußert. Harnad meint dazu, dass die Symbole, also Input und Output des Chinesischen Zimmers, und die Symbol Manipulationen auf ihrer Form und nicht auf ihrer Bedeutung basieren, sie systematisch interpretiert werden und nicht durch ihre Bedeutung, und außerdem hängt die Frage nach dem was ein Symbol System ist, von unserer Definition ab. Jedoch betont er (Vgl. Harnad, 1990): *„the interpretation will not be intrinsic to the symbol system itself, it will be parasitic on the fact that the symbols have meaning for us, in exactly the same way that the meanings of the symbols in a book are not intrinsic, but derive from the meanings in our heads”* (S. 338).

Sein eigenes Beispiel für das symbol grounding problem hat zwei Versionen, eine schwierige und eine unmögliche. Die schwierige Version besteht darin, Chinesisch als zweite Sprache zu lernen, lediglich mit einem Chinesisch/Chinesisch Wörterbuch.

Die „Reise“ durch das Wörterbuch wird mit einem „merry-go-round“ von einem bedeutungslosen Symbol zum nächsten bedeutungslosen Symbol enden, ohne bei etwas bedeutungsvollem zu halten. Jedoch ist dies kein unmögliches Unterfangen, da die Person Kenntnisse über die grundlegenden Regeln einer Sprache wie Satzbau und Grammatik verfügt. Die unmögliche Version besteht darin, Chinesisch als erste Sprache zu lernen, mit der Grundlage eines Chinesisch/Chinesisch Wörterbuches. Wobei sich Harnad die Frage stellt, *„Wie man je aus dem symbol/symbol merry-go-round aussteigen kann? und Wie die Bedeutung von Symbolen in etwas anderem als in bedeutungslosen Symbolen liegen kann?“*.

Denn die Person verfügt lediglich über eine symbolische Syntax, aus dem sich keine direkte semantische Bedeutung gewinnen lässt. Genau das stellt für Harnad das „Symbol Grounding Problem“ dar.

Ein Symbolist wie es Fodor ist würde darauf antworten, dass die Symbole ihre Bedeutung durch die richtige Verbindung des Symbol Systems mit der Welt bzw. Umwelt erhält. Harnad kontert hingegen damit, dass offensichtlich das Problem der richtigen Verbindung mit der Welt gleichzusetzen ist mit dem Problem der Kognition selbst. Denn selbst wenn jede Definition in einem Chinesisch/Chinesisch Wörterbuch auf die richtige Weise mit der Umwelt verbunden wurde, bräuchten wir trotzdem eine Definition. Weiters betont Harnad (1990):

„Unfortunately, this radically underestimates the difficulty of picking out the objects, events and states of affairs in the world that symbols refer to, i.e., it trivializes the symbol grounding problem“ (S.339).

Sein Lösungsvorschlag liegt in der Gründung eines nonsymbolic/symbolic Systems, bei dem die elementaren Symbole in zwei Arten von „nicht symbolischen“ Repräsentationen fundieren, auf diese wir anschließend noch eingehen werden.

Damit dieses System auch wirklich verstanden werden kann muss man sich ins Gedächtnis rufen wozu der Mensch in der Lage ist. Wir können unterscheiden, manipulieren, identifizieren, Objekte beschreiben, Beschreibungen erstellen und auf diese antworten. Jedoch was sind die Voraussetzungen für diese Fähigkeiten? Damit wir unterscheiden können müssen wir in der Lage sein zu urteilen, ob zwei Inputs gleich oder unterschiedlich sind. Dies wiederum basiert auf unserer Kapazität. Damit wir identifizieren können, müssen wir in der Lage sein eine Klasse von Inputs zu bilden, in der alle als äquivalent gesehen werden. Dies wiederum basiert ebenfalls auf unsere Kapazität zu entscheiden, was dazu gehört und was nicht. Um die Prozesse der Identifikation und der Unterscheidung zu verdeutlichen geben wir ein Beispiel. Wenn Sie sich ein Pferd vorstellen, dann wissen Sie dass es unterschiedliche Pferdearten gibt, wobei Sie die einen lieber und die anderen weniger mögen. Das ist die Unterscheidung. Wenn Sie ein Pferd sehen, dann werden Sie es mit ziemlicher Sicherheit auch als Pferd bezeichnen. Das ist dann die Identifizierung.

Harnad stellt sich jetzt jedoch die Frage was für eine innerliche Repräsentation benötigt wird, damit wir in der Lage sind diese zwei Fähigkeiten zu zeigen? Genau dieser Frage gehen wir im nächsten Kapitel nach.

4.5 Repräsentation

4.5.1 Ikonische Repräsentation

Ikonische Repräsentationen sind die Voraussetzung für unsere Fähigkeit, bestimmte Sachverhalte von anderen Sachverhalten zu unterscheiden oder aber Ähnlichkeiten festzustellen. *„Ikonische Repräsentationen sind geistige Bilder, die wir uns von Situationen, Gegenständen usw. machen. Diese Bilder, die keineswegs nur visuell sind, sondern auch in jedem anderen Wahrnehmungssinn verankert sein können, haben eine nachvollziehbare Ähnlichkeit mit dem von ihnen bezeichneten Objekt.“* (zitiert nach Jank, Meyer, 2002, in <http://beat.doebe.li/bibliothek/w01893.html>).

Anhand dieser ikonischen Repräsentationen sind wir in der Lage, Sachverhalte von anderen zu unterscheiden oder ihre Ähnlichkeit festzustellen. Um ein Pferd von anderen Dingen der Welt zu unterscheiden benötigen wir also ein „Bild“, eine Vorstellung von einem Pferd in unserem Kopf. Die Identifikation eines Sachverhalts ist unabhängig von der Diskrimination. Während eine Unterscheidung von Sachverhalten (wie oben beschrieben) auch dann möglich ist, wenn man den Sachverhalt nicht kennt, ist dies laut Harnad bei der Identifikation nicht möglich. Denn unsere Welt beinhaltet nicht nur Sachverhalte, die eindeutig einer Kategorie bzw. Repräsentation zugeordnet werden können, sondern es gibt auch Vermischungen, wo ikonische Repräsentationen nicht ausreichend sind, um Sachverhalte zu benennen.



Abbildung 8: Pferd oder Zebra

4.5.2 Kategoriale Repräsentation

Um einen Sachverhalt identifizieren und benennen zu können (Siehe Harnad, 1990), müssen wir unsere Vorstellung vom jeweiligen Sachverhalt auf einige wenige zentrale Merkmale reduzieren, anhand derer, Sachverhalte in bestimmte Kategorien eingeordnet werden können. Diese Einordnung, der Dinge der Welt, in Kategorien anhand zentraler (wesentlicher) Merkmale geschieht natürlich nur im Kopf und wird als „kategoriale Repräsentation“ bezeichnet.

Viele Kategorien sind uns bereits angeboren, weitere werden im Laufe des Lebens durch Erfahrung dazugelernt. So etwa die kategoriale Repräsentation vom Sachverhalt „Pferd“.

Aber: Weder die bildliche Repräsentation noch die kategoriale Repräsentation eines Sachverhalts ist symbolisch, sondern sie sind als „Kopien“, ähnlich einem Foto, der Sachverhalte, wie sie in unserer Welt vorkommen, in unseren Köpfen gespeichert, sie sind also rein sensorisch. Bei diesen Arten der Repräsentationen gibt es das Problem, wie die Vorstellungen mit der realen Welt verbunden sind, demnach nicht. Auch die Frage nach semantischen Interpretationen und deren Richtigkeit stellt sich hier nicht.

4.5.3 Symbolische Repräsentationen

Kategoriale und ikonische Repräsentationen wie oben beschrieben sind laut Harnad nicht symbolisch da sie auch nicht systematisch sind, denn dazu müsste man sie kombinieren und verändern können, wie dies bei Symbolen und Symbolsystemen der Fall ist. Allerdings sind einige Repräsentationen sehr wohl systematisch kombiniert wie zum Beispiel (Siehe Harnad, 1990):

1. *„Suppose the name „horse” is grounded by iconic and categorical representations, learned from experience, that reliably discriminate and identify horses on the basis of their sensory projections.*
2. *Suppose “stripes” is similarly grounded*
Now consider that the following category can be constituted out of these elementary categories by a symbolic description of category membership alone:
3. *“Zebra” = “Horse” & “Stripes” (S. 342).*

Die mentale Repräsentation eines Zebras ist folglich die Kombination der Symbole „Pferd“ und „Streifen“ zu einer Zeichenkette. Da „Pferd“ und „Streifen“ durch bildliche und kategoriale Repräsentationen in unseren Köpfen verankert „grounded“ sind, erbt das Zebra sozusagen die Verankerung, „zebra inherits the grounding“, und dadurch ist es eine „grounded symbolic representation“.

Rein theoretisch könnte also laut Harnad (1990) sogar jemand, der noch nie im Leben ein Zebra gesehen hat, durch verankerte Repräsentationen von „Pferd“ und „Streifen“ anhand seiner symbolischen Repräsentationen dieses identifizieren und benennen. Folglich führt die Fähigkeit, Sachverhalte der Welt anhand ikonischer und kategorialer Repräsentationen zu unterscheiden und zu benennen auch zur Fähigkeit, Sachverhalte durch symbolische Repräsentationen zu unterscheiden und zu beschreiben.

Wenn es um die Frage geht, wie die verschiedenen Repräsentationen in unseren Köpfen entstehen und wie wir die „zentralen Merkmale“ zur Kategorienbildung auswählen, kommt der oben genannte Konnektionismus wieder ins Gespräch:

Die Vorstellungen in unseren Köpfen entsteht durch konnektivistische Netzwerke, welche die Verbindung zwischen den Bezeichnungen, die wir für verschiedenste Sachverhalte haben, und den dahinterstehenden Dingen der Welt herstellen und uns helfen, die wesentlichen Merkmale, die wir zur Kategorienbildung benötigen, zu finden.

5. Lösungsansätze

Die entscheidende Frage die sich jetzt jedoch nach den obigen Erklärungen stellt lautet: Wie kann ich das „Symbol Grounding Problem“ lösen? Genau diese Antworten möchten wir im Folgenden geben.

5.1 From the Ground up

Für Harnad (1990) gibt es nur einen einzigen Weg um das Symbol Grounding Problem zu lösen, nämlich „from the ground up“. Denn ein „freies, fliegendes symbolic level“ ist kein sinnvoller, erstrebenswerter Weg, denn dies wäre gleich zu setzen mit der Entwurzelung der intrinsischen Bedeutung der Symbole.

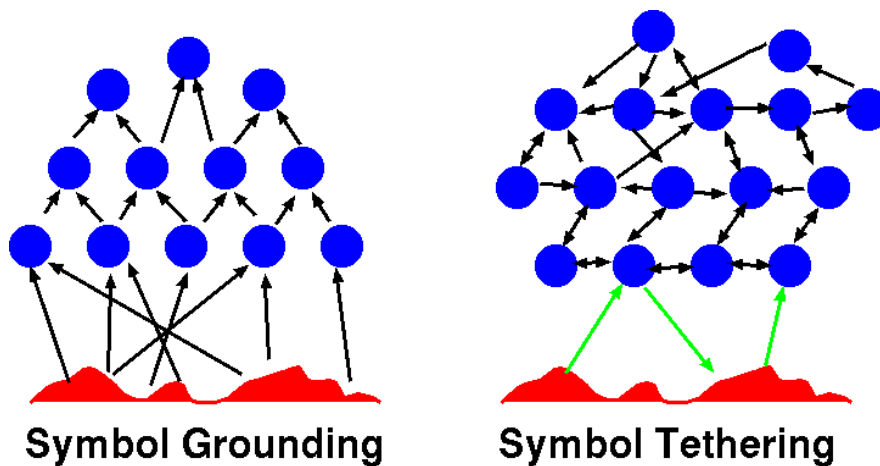


Abbildung 9: Symbolische Darstellung des Symbol Grounding Problem

In einem intrinsisch zugeordneten Symbol System bestehen mehr Grenzen bzw. Hemmungen, als in einem syntaktischen.

Harnad (1990) betont weiters: *“Symbols are manipulated not only on the basis of the arbitrary shape of their tokens, but also on the basis of the decidedly nonarbitrary “shape” of the iconic and categorical representations connected to the grounded elementary symbols out of which the higher-order symbols are composed”* (S.342).

Aufgrund dieser zwei Grenzen sind die ikonische und kategoriale Repräsentation entscheidend. Solche Symbol Systeme basieren wahrscheinlich auf kognitiven und roboterhaften Modellen, wobei Harnad betont, dass sie sich von diesen lösen sollten.

Außerdem betont er, dass: *“The present grounding scheme is still in the spirit of behaviorism in that the only tests proposed for whether a semantic interpretation will bear the semantic weight placed on it consist of one formal test and one behavioral test”* (S. 342). Sein Lösungsvorschlag lautet wie folgt:

“If both tests are passed, then the semantic interpretation of its symbols is “fixed” by the behavioral capacity of the dedicated symbol system, as exercised on the objects and states of affairs in the world to which its symbols refer; the symbol meanings are accordingly not just parasitic on the meanings in the head of the interpreter, but intrinsic to the dedicated symbol system itself” (S. 342).

Harnad betont dabei jedoch, dass dieses Modell keine Garantie darstellt, sie aber näher an ihrem Ziel sind, als gehofft.

5.2 Autonome Agenten

Der Begriff der neuen künstlichen Intelligenz wird sehr häufig mit dem Begriff „Autonome Agenten“ definiert. Autonome Agenten (Vgl. Pfeifer, Scheier, 1999) sind Roboter, Systeme, die autonom mit der Umwelt agieren, d.h. sie können sich in der Welt zurechtfinden, ohne direkt vom Menschen gesteuert zu werden. Dazu benötigen sie einen Körper, („embodied Cognition“). Diese Verkörperlichung ist zentral für ihre Intelligenz. Denn nur Akteure, die durch einen Körper in Kontakt mit der Umwelt treten können, sowie Sinnes- und Bewegungsorgane haben, können mit der Welt interagieren. Sie können die Welt wahrnehmen und in ihr handeln. Durch Sensoren sind sie mit der Umwelt verbunden. Außerdem besitzen sie bestimmte soziale Fähigkeiten, Agenten können mit anderen Agenten „kommunizieren“, auf die Umwelt und uns Menschen reagieren und außerdem zielorientiertes Handeln zeigen.

Ohne Körper ist das Symbol Grounding Problem zentral, d.h. die Roboter können keine Verbindungen zwischen ihren internen Repräsentationen und der Umwelt sowie den verschiedenen Bedeutungen der einzelnen Sachverhalte herstellen.

Grundsätzlich kann man zwei Arten von Autonomen Agenten unterscheiden, (Siehe Furbach, Obst, 2000 in <http://www.oliverobst.eu/publications/2000/FOS00.pdf>) nämlich reaktive AA und deliberative AA, wobei die reaktiven einen subsymbolischen Ansatz vertreten und mit Aktionen auf Reize reagieren können, während die deliberative Autonomen Agenten ihre Eingaben auf Symbole reduzieren und diese Symbole dann bearbeiten.

Vertreter der klassischen künstlichen Intelligenz sind der Meinung, dass diese Symbolverarbeitungsmaschinen einen Zugang zur realen Welt, also eine Verbindung zwischen Symbolen und Welt durch Mittel wie Kameraaugen, Roboterarme, usw. herstellen können. Allerdings ergeben sich auch hier einige Probleme.

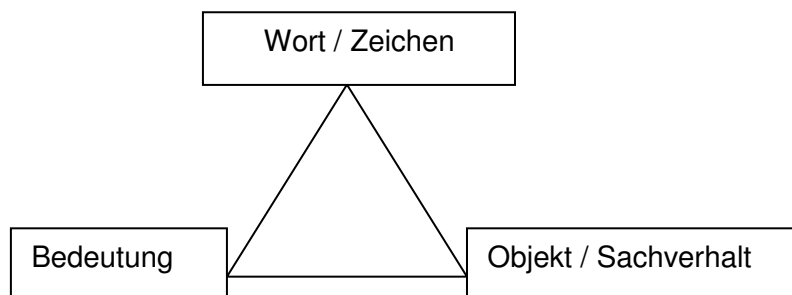
Symbol Grounding Problem

Ein KI – System arbeitet mit Zeichenketten, also aneinander gereihten Zeichen, welche das System bearbeitet.

Es erhält also eine Zeichenkette, zum Beispiel eine Anfrage, bearbeitet diese Zeichenkette, ersetzt verschiedene Zeichen und gibt schließlich wieder eine Zeichenkette aus, ohne dabei eine Verbindung zwischen verschiedenen Zeichenketten oder zwischen Zeichenketten und dem, wofür sie stehen (Sachverhalte in der realen Welt) herzustellen. Somit sind die Zeichenketten für KI – Systeme ohne jede Bedeutung.

Für einen Computer hat eine Zeichenkette wie zum Beispiel S C H W E S T E R eine bestimmte Form, die zwar bearbeitet wird, aber keinen Zusammenhang mit einer richtigen Schwester, so wie WIR unter Schwester eine weibliche Person in enger Verwandtschaft verstehen.

Das ist das Symbol Grounding Problem in der Künstlichen Intelligenz: der Computer kann keine Verbindung zwischen der Zeichenkette und dem tatsächlich und in der realen Welt dahinterstehenden Sachverhalt herstellen.



Denn um einen Sachverhalt mit einer Zeichenkette verbinden zu können, muss der Computer den Sachverhalt in der Welt identifizieren, d.h. von ähnlichen Sachverhalten abgrenzen können und eventuell Ähnlichkeiten feststellen können. Dazu braucht der Computer eine Repräsentation davon, was z.B. ein PFERD ist, was ein Pferd von anderen Vierbeinern unterscheidet, und das auch z.B. ein liegendes Pferd noch ein Pferd ist. Bei anderen Sachverhalten gestaltet sich das noch schwieriger, wenn zum Beispiel Gegenstände nicht nach dem Äußeren, sondern nach ihrer Funktion eingeteilt werden: Nicht alles was länglich ist und glänzt ist ein Messer, Messer werden viel mehr durch ihre Funktion, nämlich das Schneiden, kategorisiert und definiert (siehe Harnad, 1990).

In der klassischen Künstlichen Intelligenzforschung gibt es diesbezüglich noch weitere Probleme, auf die wir nur kurz eingehen werden. Ein zentrales Problem ist der Versuch, unsere dynamische und komplexe Welt in einem Computer zu repräsentieren. Denn die Dinge unserer Welt sind alle miteinander verknüpft und nicht selten löst eine Veränderung in der Welt zahlreiche weitere Veränderungen bzw. Nebeneffekte aus.

Ein autonomer Agent müsste nun alle potenziellen Nebeneffekte, die ein Ereignis in der realen Welt hat, überprüfen, um seine Repräsentation der Welt aktuell zu halten. Ein Beispiel dazu: Wenn ein Roboter ein Glas vom Tisch holen soll, könnten sich zahlreiche Nebeneffekte einstellen:

Das Glas könnte zerbrechen, der Tisch könnte umfallen, das Getränk im Glas könnte auslaufen und vieles mehr. Ein weiteres Problem ist das Problem der vielfältigen Umweltbedingungen. Damit eine Handlung in der realen Welt gelingen kann, müssen sehr viele Faktoren vom autonomen Agenten berücksichtigt werden. Um zum Beispiel das Glas vom Tisch zu holen, muss der Computer erst die Position des Tisches, die Form des Glases und seine Position auf dem Tisch, die Startposition des Autonomen Agenten bzw. seiner Roboterhand, eventuelle Hindernisse und zahlreiche weitere Faktoren berücksichtigen, die für uns Menschen ganz logisch sind. Außerdem müssen Autonome Agenten in Echtzeit auf der Basis unvollständiger Informationen agieren und auf unvorhergesehene Ereignisse reagieren.



Abbildung 10: Autonomer Agent

Ein Beispiel für Autonome Agenten

Roboter sind v.a. dazu da, dem Menschen das Leben zu erleichtern, schwierige und gefährliche Aufgaben zu erledigen, wie z.B. in menschenfeindlicher Umwelt wie radioaktiven Gebieten, Pipelines etc. oder aber den Menschen einfach nur zu unterhalten, wie z.B. der Roboterfußball.

AIBO

AIBO (**A**rtificial **I**ntelligence **r**o**BO**t) ist ein kleiner Roboter in Form eines Hundes, der von Sony als Unterhaltungs-Roboter und Haustierersatz entwickelt wurde. AIBO hat die Fähigkeit, von seiner Umgebung durch Versuch und Irrtum zu lernen.

So kann man ihm durch Wörter wie „good“ oder „no“ ein Feedback geben. Er ist mit Tastsensoren, Geräusch-, Stimm-, und Gesichtserkennungssensoren ausgestattet.



Abbildung 11: AIBO

6. Schlussworte

Nachdem wir uns intensiv mit dem Thema „The Symbol Grounding Problem“ beschäftigt haben, haben wir erkannt, dass die Lösungsansätze dieses Problems unzureichend und ausbaufähig sind. Damit bietet dieses Themengebiet zahlreiche weitere Forschungsmöglichkeiten.

Im Rahmen der Bearbeitung hat sich für uns nämlich die Frage gestellt ob die Autonomen Agenten wirklich eine Lösung darstellen? Denn auch die Autonomen Agenten sind auf die Unterstützung von außen, also auf eine Bedeutungszuschreibung von außen angewiesen. Dabei stellt sich jedoch die Frage ob überhaupt irgendjemand, also Mensch oder Maschine in der Lage ist die Bedeutung von Symbolen von selbst zu erzeugen. Sind wir denn nicht immer auf eine äußere, also parasitäre Bedeutungszuschreibung angewiesen? Fragen die es noch zu beantworten gilt.

Literaturverzeichnis

Duden. *Das Fremdwörterbuch*. 8. Auflage. 2005.

Fodor, Jerry A. (1975) *The Language of Thought*. New York: Crowell.

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

Pfeifer, R. & Scheier, C. (1999). *Understanding Intelligence*. Cambridge, MA: MIT Press.

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

Turing, A.M. (1950). *Computing Machinery and Intelligence*. *Mind*. New Series. Vol. 59. No. 236: 433-460.

Internetquellen:

http://en.wikipedia.org/wiki/Stevan_Harnad gefunden am 07.04.2009

http://www.techfak.unibielefeld.de/ags/wbski/lehre/digiSA/Methoden_der_KI/WS0304/vl14_a_gent.pdf gefunden am 09.04.2009

http://www2.informatik.uni-erlangen.de/Lehre/WS200304/Comp_Beauty/Folien7_4.pdf gefunden am 09.04.2009

http://support.sony-europe.com/aibo/1_1_3_aibo_story.asp?language=de gefunden am 07.04.2009

<http://de.wikipedia.org/wiki/Roboterfußball> gefunden am 13.04.2009

www.ki-team-berlin.de/pdf/robotikVortragHannover2005.pdf gefunden am 13.04.2009

www.nittka.de/download/robotik/index.html gefunden am 13.04.2009

<http://www.oliverobst.eu/publications/2000/FOS00.pdf> gefunden am 14.04.2009

www.informatik.uni-leipzig.de/~schierwa/ComputerGeistModell.pdf gefunden am 2.04.2009

<http://users.ecs.soton.ac.uk/harnad/> gefunden am 01.04.2009

<http://users.ecs.soton.ac.uk/harnad/Papers/Harnad/harnad90.sgproblem.html> gefunden am 01.04.2009

<http://de.wikipedia.org/wiki/> gefunden am 04.04.2009

http://www.uibk.ac.at/psychologie/mitarbeiter/leidlmair/the_symbol_grounding_problem.pdf gefunden am 01.04.2009

<http://de.wikipedia.org/wiki/Symbol> gefunden am 07.04.2009

<http://beat.doebe.li/bibliothek/w01893.html> gefunden am 3.04.2009

Abbildungsverzeichnis

ABBILDUNG 1: KOGNITIVE WENDE	4
ABBILDUNG 2: DARSTELLUNG VON SYMBOLEN	6
ABBILDUNG 3: BEISPIEL FÜR ZEICHEN: HIEROGLYPHEN	7
ABBILDUNG 4: SYNTAX UND SEMANTIK	7
ABBILDUNG 5: STEVAN HARNAD	9
ABBILDUNG 6: DARSTELLUNG DES TURING TESTS	9
ABBILDUNG 7: DARSTELLUNG DES CHINESISCHEN ZIMMERS	10
ABBILDUNG 8: PFERD ODER ZEBRA	13
ABBILDUNG 9: SYMBOLISCHE DARSTELLUNG DES SYMBOL GROUNDING PROBLEM.....	15
ABBILDUNG 10: AUTONOMER AGENT	18
ABBILDUNG 11: AIBO	18

Abb. 1: gefunden in <http://www.neurolabor.net/Hauptseminar%20SpracheKognition.html>

Abb. 2: gefunden in <http://www.mohawksign.com/specifications/symbols.cfm>

Abb. 3: gefunden in <http://teacher.schule.at/strassgang/seite2.htm>

Abb. 4: gefunden in <http://teacher.schule.at/strassgang/seite2.htm>

Abb. 5: gefunden in <http://www.unites.uqam.ca/cnc/images/photoharnad.jpg>

Abb. 6: gefunden in <http://www.hyperkommunikation.ch/images/turingtest.gif>

Abb. 7: gefunden in http://www.uibk.ac.at/psychologie/mitarbeiter/leidlmair/the_symbol_grounding_problem.pdf

Abb. 8: gefunden in <http://www.rp-online.de/public/bildershow/aktuelles/wissen/25033>

Abb. 9: gefunden in <http://www.cs.bham.ac.uk/~axs/fig/ground-tether.png>

Abb. 10: gefunden in http://www.iain.ira.uka.de/index_id-187_mod548tab-1.html

Abb. 11: gefunden in <http://www.neurolabor.net/Hauptseminar%20SpracheKognition.html>