

Technische Universität München

Symbol Grounding Problem

Fakultät für Informatik

Seminar Kognitive Robotik (SS12)
Sizhe Huang

Inhaltsverzeichnis

1. Intelligente Systeme
2. Formale Systeme
 - 2.1 Mentale Repräsentationen
 - 2.2 Symbol und Symbolmanipulation
 - 2.3 Symbolsystem
3. Das Symbol Grounding Problem
 - 3.1 Der Turing Test
 - 3.2 Das Chinesische Zimmer
 - 3.3 Formulierung des Symbol Grounding Problems
4. Lösungsansätze zu Symbol Grounding Problem
 - 4.1 Symbolisch und nicht-symbolisches System
 - 4.2 Embodiment
5. Kommentare zu Symbol Grounding Problem
 - 5.1 Symbolmanipulation
 - 5.2 Erkenntnisse und Modelle
 - 5.3 Genereller KI Kritik
6. Schlusswort
7. Literaturverzeichnis

1. Intelligente Systeme

Obwohl die Robotik zu einem modernen Wissenschaftszweig gehört, ist der Wunsch, intelligente Wesen künstlich zu erzeugen, uralt. Erst durch die Erfindung des digitalen Rechners wurden die Voraussetzungen für eine künstliche Intelligenz geschaffen. Die rasante Entwicklung auf dem Gebiet der Informatik erlaubt uns beispielsweise neue Spracherkennungs- und Schachprogramme zu entwickeln. Nicht selten werden sie mit dem Begriff „Intelligenz“ in Verbindung gebracht. Solche Programme erbringen zwar überdurchschnittliche Leistungen auf ihrem Gebiet, können aber aus heutiger Sicht aufgrund ihrer eingeschränkten Fähigkeiten nicht mehr als intelligent bezeichnet werden. Sie erledigen ausschließlich das, wofür sie programmiert worden sind und können niemals den vorgeschriebenen Rahmen überschreiten. In diesem Sinne ist ihr Verhalten deterministisch. Eine Maschine, die unserer Vorstellung von der Intelligenz entspricht, müsste vielmehr eigenständig denken und Aufgabe lösen können. Während einige Erfolge auf dem Gebiet der Intelligenz in den letzten Jahren erzielt werden konnten, blieb der große Durchbruch intelligente Systeme zu erschaffen dennoch aus. Es gab dementsprechend vermehrt Zweifel, ob die menschliche Denkfähigkeit überhaupt algorithmisch beschreibbar ist und nachgeahmt werden kann.

Steven Harnad, Professor für kognitive Wissenschaft an der Universität von Southampton, leistete seinen Beitrag für die Forschung der künstlichen Intelligenz, als er im Jahr 1990 das Symbol Grounding Problem veröffentlichte. Darin diskutiert er den Einfluss der kognitiven Forschung in den vergangenen Jahren auf die Entwicklung der künstlichen Intelligenz. Außerdem stellte er sich die entscheidende Frage, wie Maschinen an die Bedeutung der Symbole gelangen können und wie Symbole mit der Umwelt in Verbindung gebracht werden könnten.

Dieser Artikel versucht einen kompakten Überblick über das Symbol Grounding Problem zu geben. Zunächst werden die Symbolsysteme behandelt, sie dienen als Grundlage für die spätere Problemformulierung. Anschließend wird das Symbol Grounding Problem mithilfe des chinesischen Zimmers erläutert.

2. Formale Systeme

2.1 Mentale Repräsentationen

Im Gegensatz zum direkten Realismus legt die Theorie der mentalen Repräsentation zugrunde, dass der Mensch seine Reize und Wahrnehmungen, die er durch seine Umwelt erfährt, in Form von Repräsentationen abspeichert. Ein Phänomen aus der Realität wird zunächst abstrahiert und dann durch ein Symbol in das kognitive System übersetzt. Die Repräsentation eines Objekts stellt somit die Verbindung zwischen unserem Gehirn und einem realen Objekt dar. Symbolische Informationen können so in unserem Gehirn verarbeitet und gespeichert werden. Sie bilden die Basis für die Strukturierung der gesammelten Erfahrung.

Von einer intelligenten Maschine wird deshalb eine kognitive Leistung gefordert, die der Menschlichen gleichwertig ist. Um sich Zugang zur Realität zu verschaffen, muss eine Maschine folglich ebenfalls mithilfe der Repräsentationstechnik arbeiten können. Die Bedeutung des Symbols ist daher sowohl für Menschen als auch für Maschinen elementar.

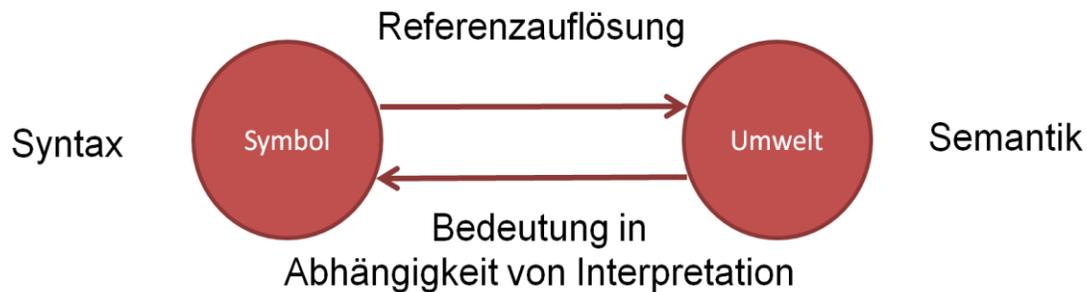
2.2 Symbol und Symbolmanipulation

Für Harnad sind Symbole eine Reihe von willkürlich gewählten Zeichen. Er unterscheidet generell zwei Arten von Symbolen: primitiven Symbolen und zusammengesetzte Symbole, die wiederum aus primitiven Symbolen bestehen. Symbole sollten Raum für semantische Interpretationen bieten. Sie sollten durch explizite Regeln verändert bzw. manipuliert werden können, wodurch neue Symbole entstehen. Symbolmanipulation ist also ein Vorgang, bei dem eine explizite Regel auf ein Symbol angewendet wird, um dessen Darstellung bzw. Ordnung zu verändern. Wichtig ist dabei, dass die Manipulation ausschließlich auf der syntaktischen und nicht auf der semantischen Ebene stattfindet. Die Sinnhaftigkeit eines Zeichens, die von ihrer Syntax unabhängig ist, bleibt daher von der Manipulation unbeeinflusst.

2.3 Symbolsystem

Ein Symbolsystem besteht aus einer endlichen Menge von Symbolen und expliziten Manipulationsregeln, die selbst wiederum durch Symbole formuliert werden können. Die Beziehung zwischen einem Zeichen als Referenz und dem referenzierten Objekt hängt lediglich von dem manipulierenden System ab.

Das Wort „Wurzel“ würde bei einem Botaniker, einem Mathematiker und einem Zahnarzt wahrscheinlich verschiedene Vorstellung hervorrufen. Das liegt daran, weil ein Wort unterschiedliche Bedeutungen besitzen kann, die aber erst interpretiert werden müssen.



Mit Übung können wir Menschen aus dem Kontext erkennen, welche der Bedeutung wirklich gemeint ist. Aber woher soll eine Maschine die Regel kennen, aus der wir die Bedeutung eines Wortes ableiten? Harnad weist daraufhin, dass der Sinn eines Symbols nicht instinktiv, sondern parasitär ist. Es ist nicht möglich mithilfe ihrer Syntax auf die Semantik des Wortes zu schließen, da sie, wie schon erwähnt, unabhängig voneinander existieren. Die Symbole tragen die Regel, wie sie interpretiert werden sollen, nicht bei sich. Diese Regeln müssen viel mehr von außen hinzugefügt werden. Alleinstehende Symbole ohne Bezugssysteme tragen daher keine Bedeutung. Erst durch die Verknüpfung mit einem Symbolsystem erhält ein Zeichen ihre eindeutige Sinnhaftigkeit.

Unsere natürliche Sprache kann als Beispiel für ein formales System gesehen werden. Die Buchstaben sind die primitiven und atomaren Symbole. Die Syntax eines Zeichens kann dabei willkürlich gewählt werden, da sie lediglich eine gesellschaftliche Vereinbarung sind. Aus den Alphabeten kann man zusammengesetzte Zeichenketten erstellen, die Wörter und Sätze bilden. Die Grammatik und die Rechtschreibung sind in diesem Fall Vorschriften, wie man Zeichen und Symbole aufgrund ihrer Syntax manipulieren darf.

3. Das Symbol Grounding Problem

3.1 Der Turing Test

Im Jahr 1950 schlug der damals angesehene Mathematiker und Pionier der Informatik Alan Turing den folgenden nach ihm benannten Turing Test vor, um die Intelligenz einer Maschine zu bestimmen. Der Ausgangspunkt für ihn war die zentrale Frage: „Können Maschinen denken?“

Eine Testperson soll ein Gespräch mit zwei ihr unbekanntem Gesprächspartnern führen, wobei einer von ihnen ein Mensch und der andere eine Maschine ist. Das Ziel der Testperson ist durch Fragestellung herauszufinden, bei welchem der beiden Gesprächspartner es sich um eine Maschine handelt. Die getestete Person hat weder direkt Hör- noch Sichtkontakt mit ihnen. Die einzige Kommunikationsmöglichkeit ist ein Gespräch auf einem zwischen ihnen verbundenen Monitor zu führen. Nach Turing könnte ein Dialog wie folgt aussehen: ¹

Q: Please write me a sonnet on the subject of the Forth Bridge.

A: Count me out on this one. I never could write poetry.

Q: Add 34957 to 70764.

A: (Pause about 30 seconds and then give as answer) 105621.

Q: Do you play chess?

A: Yes.

Wäre die Person nach dem Gespräch trotzdem nicht in der Lage die Maschine eindeutig zuzuordnen, so hat die Maschine den Turing Test bestanden. Solch ein System kann folglich als intelligent bezeichnet werden, weil das Testsubjekt von der Imitation getäuscht wird.

Der Turing Test sorgte in einigen Wissenschaftsbereich nicht nur für Aufsehen sondern auch für Kritik. Ein genereller Kritikpunkt war, dass eine Unterhaltung

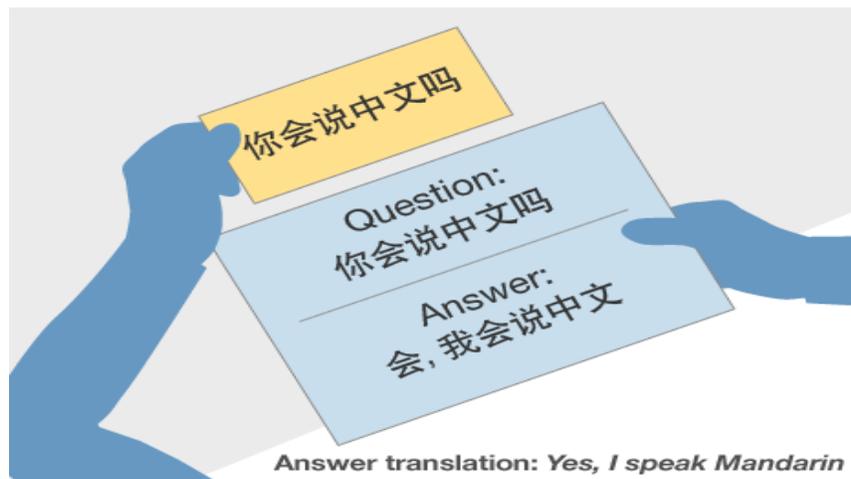
¹ Vgl. <http://www.loebner.net/Prizef/TuringArticle.html>, Computing Machinery And Intelligence

allein wenig über das Denkvermögen eines Wesens aussagt. Demnach kann ein Wesen weiterhin intelligent sein, ohne dass es in der Lage ist, mit Menschen kommunizieren zu können. Es ist weiterhin zweifelhaft, ob der TT ein ausreichendes Kriterium für die Intelligenz ist, da die Messung von Intelligenz in diesem Fall ausschließlich aus menschlicher Sicht geschieht.

3.2 Das Chinesische Zimmer

Das in der Fachliteratur weitaus bekanntere Gegenargument zum Turing Test ist das 1980 von Searle R. John entwickelte Chinesisches Zimmer. Es soll zeigen, dass die Kriterien aus dem Turing Test unzureichend sind, um über den Intellekt einer Maschine urteilen zu können. Die These von Searle lautet wie folgt:

In einem geschlossenen Raum befindet sich eine Person, die kein Chinesisch versteht. Durch die Türspalte bekommt sie einen Fragebogen auf Chinesisch, die sie beantworten soll. Anschließend erhält sie ein ebenfalls in Chinesisch geschriebenes Handbuch als einziges Hilfsmittel, worin die Lösungen bereits auf die Fragen abgebildet werden. Es würde ihr ermöglichen alle Fragen auf dem Zettel zu beantworten.



(Quelle: <http://www.bbc.co.uk/news/technology-17547694>)

Obwohl die Testperson kein chinesisches versteht, wäre sie allein mithilfe des Handbuches in der Lage, den Fragebogen zu beantworten. Sie müsste dazu nur die Zeichen aus dem Fragebogen mit den Zeichen aus dem Handbuch bezüglich

ihrer Darstellung vergleichen. Danach könnte sie die Lösungen aus dem Handbuch ablesen und abschreiben, ohne sie dabei verstehen zu müssen.

Eine weitere Person, die sich außerhalb des Raumes befindet und der chinesische Sprache mächtig ist, bekommt anschließend das Antwortblatt der sich im Zimmer befindenden Person. Sie würde daraufhin irrtümlich annehmen, dass eine chinesisch verstehende Person sich im Raum befindet, weil das Antwortblatt korrekt beantwortet wurde.

Mit dieser Analogie möchte Searle zeigen, dass Maschinen, die Symbole lediglich syntaktisch verarbeiten keineswegs als intelligent bezeichnet werden können. Wesen, die rein formal Algorithmen ausführen, begreifen auf diese Weise nicht die Sinnhaftigkeit ihrer Tätigkeit. Im Gegensatz zu Turing ist Searle deshalb der Meinung, dass Intentionalität eine notwendige Voraussetzung für die Intelligenz ist. Der Turing Test kann deshalb nach ihm nicht als Maßstab für die Intelligenz gesetzt werden. Wenn wir solche Maschinen als intelligent bezeichnen, dann liegt das einzig daran, weil wir ihnen aufgrund unserer Vorstellung des Ergebnisses Denkvermögen zusprechen.

3.3 Formulierung des Symbol Grounding Problems

Steven Harnad erläutert das Symbol Grounding Problem, indem er eine schwierige und eine unmögliche Version des chinesischen Zimmers vorstellt.

Die erste Version besteht darin chinesisch als zweite Sprache zu lernen, wobei das einzige Hilfsmittel ein chinesisch/chinesisches Wörterbuch ist. Um herauszufinden, was die Bedeutung eines Wortes ist, müsste man die Erklärung des Wortes in dem Wörterbuch verstehen, die selbst wieder aus Wörtern besteht. Die Suche in dem Wörterbuch würde in eine Endlosschleife enden, da die Semantik eines Wortes immer in den Bedeutungen andere Wörter liegen. Um überhaupt den Einstieg zu schaffen, bräuchte man bestimmtes Wissen, das a priori bekannt ist. Das Problem kann nach Harnad aber dennoch gelöst werden, weil die chinesisch lernende Person bereits gewisse Vorkenntnisse über Sprachen besitzt. Sie verfügt also über ein verankertes Symbolsystem und Erfahrung aus der Umwelt. Mithilfe dieser als Basis, wäre sie in der Lage chinesisch auf sehr mühsame Weise zu lernen.

Die unmögliche Version des chinesischen Zimmers beschreibt die Schwierigkeit, Chinesisch als erste Sprache zu lernen, wobei das einzige Hilfsmittel wieder nur ein chinesisches/chinesisches Wörterbuch ist.

Dieser Sachverhalt ist für Harnad das Symbol Grounding Problem. Ein rein formales Symbol, das in einem System verankert ist, bekommt seine Bedeutung nur durch andere formale Symbole, die wiederum durch andere Symbole beschrieben werden müssen. Die Bedeutung eines Symbols aus einem rein symbolischen System ist für Maschinen somit nicht zugänglich. Wörter sind allein in unseren Köpfen verankert. Wäre ein Wort auf einem Papier oder einer Festplatte, so ist es nicht verankert. Anders ausgedrückt: das Symbol Grounding Problem handelt von der Schwierigkeit, verschiedene Symbole mit ihren entsprechenden Bedeutungen aus der Umwelt zu verknüpfen, ohne auf externe Interpretationsprozesse zurückgreifen zu müssen.

4. Lösungsansätze zu Symbol Grounding Problem

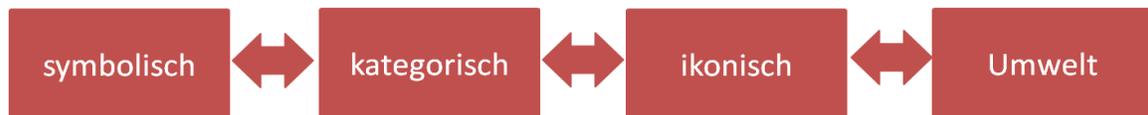
4.1 Symbolisch und nicht-symbolisches System

Harnad schlägt daraufhin ein symbolisches und nicht-symbolisches System vor, um sich dem Problem anzunähern. Zunächst geht er davon aus, dass Menschen bestimmte kognitive Fähigkeiten besitzen müssen, da wir in unsere Umwelt sonst nicht zurechtkommen würden. Dazu zählt er die Fähigkeit zu diskriminieren, identifizieren und Objekte zu beschreiben.

Ikonische Repräsentation sind „Bilder“, die wir von unserer Umwelt machen. Sie sind interne Projektion eines Objekts aus der Umwelt auf unserem sensorischen Träger. Auf dieser Ebene werden Sachverhalte durch visuelle Bilder dargestellt. Anhand dieser ikonischen Repräsentationen sind wir in der Lage, verschiedene Sachverhalte voneinander zu unterscheiden und zu diskriminieren. Wir vergleichen die Ähnlichkeit einer ikonischen Repräsentation mit der eines anderen und urteilen schließlich darüber, ob sie gleich oder ungleich sind.

Eine andere wichtige Fähigkeit des Menschen liegt darin, Objekte zu identifizieren. Hierbei ist die Diskriminierung unabhängig von der Identifikation. Man wäre zum Beispiel in der Lage Objekte zu diskriminieren, ohne das Objekt

genau zu kennen. Infolgedessen ist Harnad der Meinung, dass die ikonische Repräsentation allein nicht ausreicht, um unsere Fähigkeit der Identifikation zu begründen. Es wäre nicht möglich die Vielzahl der ähnlichen Objekte in der Welt allein mithilfe der ikonischen Repräsentation zu unterscheiden. Der Übergang von einem Objekt in das Nächste könnte flüssig sein. Die Ikonen der Sensoren sind dazu nicht selektiv genug. Um Identifizieren zu können, müssen Ikonen auf ihre Invariante reduzierbar sein. Erst dann können Objekte einer Kategorie, die diese Eigenschaft besitzt, von Objekten anderer Kategorien unterschieden werden. Wichtig ist dabei, dass sowohl die ikonische als auch die kategorische Repräsentation nicht symbolisch sind. Das Erstere ist eine analoge Kopie der sensorischen Projektion der Umwelt, das Letztere filtert einige markante Merkmale heraus, sodass man sie anschließend in Kategorien zuordnen kann. Da sie beide auf sensorische Ebene arbeiten, gibt es keine Probleme bei der Referenzauflösung. Ihre Verbindung mit der wirklichen Welt ist eindeutig. Zuletzt sei bemerkt, dass sie keine Bedeutung besitzen und es darüber hinaus keinen Sinn macht, von ihrer Bedeutung zu sprechen. Analog hat ein Bild, das von einer Kamera gemacht wird, keine Bedeutung, da es nur seine Umgebung auf einem Papier projiziert.



Um kategorische Repräsentationen manipulieren zu können, müssen diese in ein System übertragen werden. Wir wissen bereits, dass neue Bedeutungen durch Kombination der symbolischen Repräsentationen entstehen können. Anders als in einem rein symbolischen System können die symbolischen Repräsentationen auf ikonische und kategorische Repräsentation zurückgreifen. Die Symbole können somit in der ikonischen und kategorischen Repräsentation verankert werden. Dazu existiert das folgende Beispiel von Steven Harnad selbst:

Es wird angenommen, dass die Wörter „Pferd“ und „Streifen“ in der ikonischen und kategorischen Repräsentation bereits verankert sind. Nun kann das Zeichen „Zebra“ = „Pferd“ + „Streifen“ als formales Symbol generiert werden. Sie ist jetzt eine symbolische Beschreibung der Kategorie Zebra durch andere formale Symbole. Das Zeichen „Zebra“ repräsentiert also nichts anderes als „Pferd“ + „Streifen“. Da „Pferd“ und „Streifen“ ikonisch und kategorisch verankert sind, ist ebenso das neu erzeugte Wort „Zebra“ verankert. Harnad glaubt, dass jemand,

der lediglich die Bedeutung der Wörter „Pferd“ und „Streifen“ kennt, in der Lage wäre, das neue Symbol „Zebra“ mit einem Objekt aus der realen Welt zu identifizieren.

4.2 Embodiment

Ein virtueller Flieger fliegt nicht, eine virtuelle Kochplatte kocht nicht. Es sind nur Symbole, die auf unsere gedankliche Interpretation zurückgehen. Nach dieser Argumentation kann ein virtuelles System auch nicht denken. Folglich können sie nicht intelligent sein. Sie brauchen vielmehr einen physischen Körper, damit sie direkt Verbindung mit ihrer Umwelt aufnehmen können. Erst wenn sie in der realen Welt eingebettet sind, können sie die Umgebung durch ihre Sensoren wahrnehmen. Die Bedeutung eines Zeichens wäre dann in den Sensoren, die ein System wahrnimmt, direkt verankert. Eine Beschreibung der Symbole durch andere Symbole wäre daher nicht mehr notwendig und auch überflüssig, da ein System die Bedeutung selbstständig durch die Umgebung bilden kann.

Im Gegensatz zur klassischen Auffassung von Intelligenz wird die These, dass Intelligenz einen physikalischen Körper voraussetzt, in der Kognitionswissenschaft als eine grundlegende Wende angesehen. Sie findet in verwandte Wissenschaftsbereiche ebenfalls Anerkennung. Beispielsweise wird dieser Ansatz in der Psychologie verwendet, um die Wechselwirkung zwischen Körper und Kognition zu betonen.

5. Kommentare zu SGP

Seit der Veröffentlichung existieren verschiedene Meinungen und Kommentare zu dem Problem. Während manche Wissenschaftler glauben, das Problem schon gelöst zu haben, behaupten andere wiederum, das Problem sei überflüssig und trage nicht zu der KI Entwicklung bei.²

² Vgl. <http://homepage.univie.ac.at/nicole.rossmanith/concepts/papers/steels2008symbol.pdf>

5.1 Symbolmanipulation

Bei der bisherigen Argumentation haben wir angenommen, dass die Symbolmanipulation kein ausreichendes Kriterium für die Intelligenz ist. Was aber wenn die Annahme falsch ist und Symbolmanipulation doch ein ausreichendes Kriterium für die Intelligenz ist? Wesen, die Symbole manipulieren, sind auf ihre eigene Art und Weise „intelligent“. Das Symbolmanipulierende Subjekt muss nicht unbedingt die Bedeutung der Tätigkeit verstehen, wenn seine Tätigkeit an sich für ein größeres System einen Sinn ergibt.

5.2 Erkenntnisse und Modelle

Man kann sich die Frage stellen, ob Symbole wirklich in unseren Köpfen existieren, da das Gehirn doch nur aus Neuronen und Synapsen besteht. Alternativ könnte man Wörter als akustische Reize sehen. Sie sind in dem Sinne ein „Symbol“, weil sie bestimmte Ereignisse in unserem Gehirn hervorrufen. Andere Wissenschaftler gehen davon aus, dass unser Gehirn nicht direkt mit Symbolen arbeitet sondern durch numerische Vektoren erklärt werden kann³. Wie gut sind unsere Modelle über Intelligenz wirklich? Reicht unser bisheriges Wissen aus der Neurowissenschaft aus, um das SGP formulieren und verstehen zu können?

5.3 Genereller KI Kritik

Hubert Dreyfus kritisiert die grundlegenden Annahmen der früheren KI Forschung. Die Annahme, dass das gesamte Wissen formalisiert werden könnte, wird dabei scharf kritisiert. Diese Kritik betrifft insbesondere auch das SGP. Er argumentiert dass viel von unserem menschlichem Wissen nicht symbolisch ist und deshalb nicht explizit für Maschinen verständlich gemacht werden kann. Als Beispiel nennt er eingeübte Fähigkeiten oder Alltagswissen. Es ist daher nicht möglich alles, was für intelligentes Verhalten wichtig ist, explizit darzustellen oder zu formalisieren. Folglich ist der Versuch, intelligente Maschinen auf dem bisherigen Weg zu erzeugen, zum Scheitern verurteilt.

³ Vgl. Solutions and Open Challenges for the Symbol Grounding Problem, Seite 72

6. Schlusswort

Das SGP bleibt ein aktuelles Thema in der KI Entwicklung. Es ist einerseits interessant, weil viele Themen aus verwandten Wissenschaftsbereichen der KI darin involviert sind. Andererseits machen die vielfältigen Betrachtungsmöglichkeiten das Problem schwer greifbar. Es erfordert umso mehr einen zusammenhängenden Überblick und Fachwissen aus mehreren Bereichen, um das Problem gründlich behandeln zu können.

Die Lösungsansätze, die hier behandelt wurden sind gewiss nicht die einzigen Versuche das Symbol Grounding Problem zu beheben. Es ist nach wie vor umstritten, ob das Symbol Grounding Problem gelöst ist. Sicher ist, dass die Bedeutung von SGP für die Entwicklung der KI sich erst in der Zukunft zeigen wird.

7. Literaturverzeichnis

http://www.cs.ox.ac.uk/activities/ieg/research_reports/ieg_rr210605.pdf

http://www.scholarpedia.org/article/Talk:Symbol_grounding_problem

http://www.scholarpedia.org/article/Symbol_grounding

<http://plato.stanford.edu/entries/computational-mind/>

<http://www.loebner.net/Prizef/TuringArticle.html>

<http://www.tech.plym.ac.uk/SoCCE/ITALK/publications/Cangelosi-IJSS-2011.pdf>

<http://courses.media.mit.edu/2004spring/mas966/Harnad%20symbol%20grounding.pdf>

<http://www.cse.unsw.edu.au/~achim/Research/Philosophie/node13.html>

<http://cogprints.org/2132/3/harnad01.cangelosi-greco.pdf>

http://www.scholarpedia.org/article/Chinese_room_argument

<http://www.cognitivephilosophy.net/consciousness/human-cognition-and-the-chinese-room/>

<http://www.bbc.co.uk/news/technology-17547694>

<http://homepage.univie.ac.at/nicole.rossmanith/concepts/papers/steels2008symbol.pdf>