

# Masterarbeit

am

Lehrstuhl für Kommunikationstechnik

Prof. Dr.-Ing. habil. Matthias Wolff

Brandenburgische Technische Universität Cottbus - Senftenberg  
Fakultät für Maschinenbau, Elektrotechnik und  
Wirtschaftsingenieurwesen  
Institut für Elektronik und Informationstechnik

## **Natur- und geisteswissenschaftliche Grundlagen kognitiver Systeme**

Eingereicht von:

Name: Daniel Götze

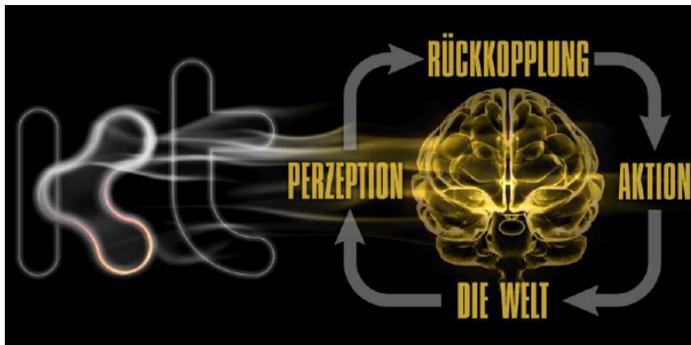
Matrikel: 2831780

Anschrift: Räschener Str.18  
03048 Cottbus

Betreut von:

Dr.-Ing. Ronald Römer

Cottbus der 11.11.2016



Thema:

## Natur- und geisteswissenschaftliche Grundlagen Kognitiver Systeme.

Beschreibung:

Für die Beschreibung technischer kognitiver Systeme wird naturgemäß auf biologisches oder psychologisches Begriffsinventar zurückgegriffen. Eine allgemein akzeptierte Definition für Kognition oder kognitive Systeme findet man in den Ingenieurwissenschaften jedoch nicht in ausreichendem Maße. Diese Situation kann vor allem durch den hohen Grad an Interdisziplinarität auf diesem Forschungsgebiet begründet werden. Für eine ingenieurwissenschaftliche Betrachtung ist es jedoch unbedingt erforderlich, den Geltungsbereich technischer kognitiver Systeme zu definieren. Insbesondere eine Abgrenzung zur Kybernetik sollte herausgearbeitet werden. Für diese Arbeit kann man sich - bei Bedarf - an den Vorarbeiten des Betreuers orientieren.

Folgende Teilaufgaben sind zu bearbeiten:

- Darstellung kognitiver Systeme unter Einnahme
  - einer biologischen, psychologischen und einer philosophischen Perspektive.
- Erschließung von Brückenkonzepten, mit denen zwischen den einzelnen Disziplinen vermittelt werden kann
  - Konzepte die für Organismen und Automaten gleichermaßen gelten, so dass ein Vergleich von organischen- und künstlichen Systemen gerechtfertigt ist.

Betreuer:

Dr.-Ing. Ronald Römer

T: 0355 695007

E: ronald.roemer@tu-cottbus.de

## Kurzfassung

Diese Arbeit hat das Ziel sowohl die Natur- als auch die Geisteswissenschaftlichen Grundlagen, die für kognitive Systeme notwendig sind zusammenzutragen.

Dabei werden verschiedene Aspekte der Verhaltenspsychologie betrachtet um Grundcharakteristika für technische Systeme zu ermitteln. Diese Charakteristika werden hierbei mathematisch formuliert und in einer Simulationssoftware genutzt werden um ein lernendes System zu implementieren.

Dabei wird auf das von Shannon entworfene Prinzip der „Theseuschen Maus“ eingegangen, eine technisch entworfene Maus, die das Labyrinth in dem sie sich befindet erlernt und so immer einen geeigneten Weg zum Ziel findet.

Zu Grunde liegt dabei die von M. Sc. Werner Meyer angefertigte Masterthesis, die sich mit dem Prinzip des Lernens beschäftigt, das in MATLAB simuliert umgesetzt wurde.

Als Erweiterung der Thematik wurde diesem Programm ein Coping Algorithmus hinzugefügt, welchem dem Agenten ermöglicht sich auch in einer Umgebung, die sich verändert zu lernen und sich so anzupassen.

Weiterhin wurde eine Visualisierung des Coping Prozesses in Form einer Merkmals-Werte-Relation implementiert, sodass es möglich ist zu sehen welche Aktionen der Agent nun zur Verfügung hat.

## Abstract

This disquisition shall assemble the basics of natural- and humanistic fundamentals, which are necessary for cognitive systems. Different aspects of behavioral psychology are considered to get the basic character for implement technical systems. These psychologic characteristics get formulate in mathematic equations and implement in a specific simulation software to get an adaptive system.

The basic for this system is the "Theseus Mouse", developed by Claude Shannon.

This mouse is learning the hole maze and can find the best way to reach his aim.

This disquisition continues the master thesis from M. Sc. Werner Meyer, which treats with psychologic basics about learning. These basics are implemented in MATLAB to create an adaptive System. Furthermore, a new coping algorithm is described, which allows the cognitive agent to create new actions in a modifiable environment.

Additionally, a visualization was created, that represents the coping process in a feature value relation, which shows every possible action.

# Inhalt

Kurzfassung.....	3
Abstract .....	4
1. Einleitung .....	3
2. Richtungen der Psychologie .....	5
2.1 Behaviorismus .....	5
2.2 Konstruktivismus/Introspektion .....	6
2.3 Biologische Perspektive .....	9
2.4 Finale Systeme.....	12
3. Technische Modellierung kognitiver Systeme.....	14
3.1 Der kognitive Kreis .....	14
3.2 Lernformen .....	17
3.2.1 Biologisch orientiertes Lernen.....	17
3.2.2 Formalisierung des Reinforcement Learning.....	19
3.2.3 Adaption und Coping.....	22
3.3 Leistungsfähigkeit.....	25
4. Kritik am traditionellem kognitivem Kreis .....	27
5. Integration des Coping Algorithmus.....	30
6. Experimentiersystem: kognitive Maus „Theseus“ .....	32
6.1 Der kognitive Kreis an der Theseuschen Maus.....	32
6.2 Messung des Lernaufwands .....	33
6.2 Implementierung des Coping .....	36
6.4 Visualisierung der Merkmal-Werte-Relation .....	39
7. Fazit .....	43
Literaturverzeichnis .....	i
Abbildungsverzeichnis.....	ii
Eidesstattliche Erklärung .....	iii

## 1. Einleitung

Die Welt von heute ist eine Welt der Veränderung, eine Welt in der der technische Fortschritt sich seinen Weg bahnt. Immer neue Produkte kommen in Sekundentakt auf den Markt und verdrängen Altbewährtes und Gewohntes. Der Mensch schafft sich eine immer komfortablere Welt.

Wir Menschen sind schon immer darauf aus, unsere Aufgaben schnell und effektiv zu erledigen und dennoch ein höchstes Maß an Sicherheit an den Tag zu legen. Angefangen damit, wie man das Jagen und Kämpfen optimieren kann, hat der Mensch Waffen entwickelt mit denen er Gefahren für sich minimieren kann. Dies ging dann über auf die Landwirtschaft, in dem man Tiere domestizierte und sich das Jagen so ersparen konnte. Dieses Bestreben alles angenehmer und einfacher zu machen hält noch bis heute an. Selbst heute will der Mensch Kleinigkeiten und Nebensächlichkeiten im Alltag effektiv und schnell erledigt haben, seine Arbeit so schnell wie möglich erledigt haben, damit er sein Augenmerk auf seine Hauptprobleme legen kann.

Dazu benötigt man aber Systeme, die sich dem Benutzer anpassen, vom ihm lernen und sich so verhalten wie es gewünscht wird. Früher war so etwas in Form eines Dieners dem Adel vorbehalten, aber da diese auch Menschen waren hatten sie eigene Bedürfnisse und waren Fehleranfällig und meist nicht so effektiv wie man sich das wünschen würde. Eine technische Lösung, die einen „Diener“ für uns ermöglicht wäre ein Traum für jeden Haushalt und jeden Menschen.

Diese Systeme werden geführt unter der Bezeichnung kognitive Systeme. Kognition ist ein Begriff der in der Psychologie und Wissenschaft unterschiedlich gehandhabt wird. In der Psychologie bezeichnet Kognition mentale Prozesse wie Wünschen, Gedanken und Meinungen. Die Technik beschränkt sich eher auf die praktischen Teile und bezeichnet Kognition als eine Informationsverarbeitung. Damit ist gemeint, dass ein eingehender Reiz in die Kognition eines Individuums eintritt und eine Handlung auslöst.

In diesem Teil der kognitiven Psychologie wird von keinem bestimmten Lebewesen ausgegangen. Es werden immer allgemeine Fakten betrachtet und aufgezeigt, dass dies bei anderen Lebewesen in der Natur nachweisbar ist.

So unterscheidet man zwischen der Seele, und der Psyche. Die Psyche ist der Teil, der in Partnerschaft mit dem Körper wirkt, die Verbindung von Körper zur Seele. Die Seele ist der Teil, der den Menschen und seine geistigen Fähigkeiten ausmacht, wie Wünsche, Gedanken, Emotionen etc.[BIS09]

Der Teil der für die Technik selber interessant ist, ist die Psyche. Eine Maschine zu entwickeln die eine eigene Seele hat, also über Emotionen und einen eigenen Willen verfügt ist nach dem deutschen Psychologen Dietrich Dörner möglich, würde aber ethische Fragen aufwerfen über die Verwendung und den Umgang mit Maschinen.

Jegliche Signale die von außen auf das Subjekt einwirken werden als Phänomen bezeichnet. All diese Signale werden vom Gehirn interpretiert und bilden eine Vorstellung der Umgebung anhand der empfangenen Signale und der vorhandenen Erfahrungen. Diese Ebene der Psyche wird als phänomenale Ebene bezeichnet. Wenn z.B. ein Tier das Brüllen eines anderen Tieres hört, nimmt er dies auf und erkennt anhand seiner Erfahrungen wie groß die Bedrohung für ihn ist. Es könnte ein überlegenes Tier, ein Artgenosse oder eine Beute sein.

Thema dieser Arbeit wird es sein, die natur- und geisteswissenschaftlichen Grundlagen von Kognition zu erfassen um beschreiben zu können, wie ein lernfähiges System implementiert werden kann.

## 2. Richtungen der Psychologie

### 2.1 Behaviorismus

Der Behaviorismus leitet sich vom englischen Wort „behavior“ ab, was übersetzt Verhalten bedeutet. Diese Bezeichnung ist nahezu selbsterklärend, denn bei dieser Richtung der Psychologie geht es darum das Verhalten des Subjektes auf gewisse Reize zu untersuchen. Der Begriff selber wurde 1913 in einer Facharbeit von John Broadus Watson, einem US-amerikanischen Psychologen, geprägt. [WAL04]

Einer der bekanntesten Vertreter ist Iwan Petrowitsch Pawlow, der mit seinen Experimenten zur klassischen Konditionierung bekannt wurde. Diese war die Grundlage zum gezielten Abrichten von Tieren oder zur Desensibilisierung von z.B. Phobien.

Pawlow wendete dieses Prinzip bei seinem Hund an. Er stellte fest, dass sein Hund, wenn er mitbekommt, dass er Fressen bekommt, automatisch mehr Speichel bildet. Bei jeder Mahlzeit läutete er nun eine Glocke. Nach einer Weile stellte er fest, dass er nur noch die Glocke läuten musste und sein Hund automatisch anfang Speichel zu bilden, da er Essen erwartete, obwohl keines da war.[GES04]

Der Behaviorismus ist die Grundlage jeder Verhaltensforschung von Psychologen. Es geht dabei darum, auf einen unbedingten Reiz zu erkennen, welche unbedingte Reaktion noch folgt. Angstpatienten, die z.B. unter Höhenangst leiden, unterliegen der Reaktion Angst, sobald sie einer größeren Höhe ausgesetzt sind. Setzt man diese Patienten jetzt in anfangs schwacher Form dem angstausslösenden Reiz aus und lässt sie dabei einer Aktivität, die sie beruhigt, z.B. Entspannungsübungen, nachgehen, werden sie desensibilisiert für den Angstreflex und können ihn auf Dauer so verlieren.

Anders als bei der Introspektion, bei der nur das eigene Erleben des Subjektes betrachtet wird, verlangt der Behaviorismus keine Betrachtung der inneren

Vorgänge und Empfindungen, was der größte Kritikpunkt am Behaviorismus darstellt. Das Gehirn wird wie eine Black Box behandelt, in dem man auf naturwissenschaftliche Weise nur beobachtet, welche Reaktion es tätigt bei einem bestimmten Reiz. Jedoch wird bei dieser Sichtweise das Prinzip des Motivs vernachlässigt. Hält man einem Hund was zu essen hin kann man beispielsweise beobachten, dass er das eine mal weggeht, das andere Mal das Essen nimmt. Dieses Verhalten ist vom inneren Zustand des Hundes, nämlich Hunger, abhängig, was vom Behaviorismus allerdings nicht beachtet wird.

Watson experimentierte ähnlich wie Pawlow mit Reflexen und stellte die Theorie auf, dass Organismen nur durch auf sie wirkende physikalische Reize etwas über ihre Umgebung in Erfahrung bringen könnten. Jede Reaktion des Organismus sei umweltbedingt. Um das Problem der inneren Zustände zu umgehen, definierte er diese (Magenknurren, volle Blase etc.) auch als Umwelt. Dies berücksichtigt aber nicht, dass diese Zustände unter Umständen abhängig vom Subjekt sind und nicht zufällige Reize. [BIS09]

## 2.2 Konstruktivismus/Introspektion

Die Introspektion war zu Beginn des 20. Jahrhunderts die vielversprechendste Methode zur Erforschung der menschlichen Psyche. Forscher wie Franz Brentano und Edward Bradford Titchener verschrieben sich ihr um den Geist des Menschen erforschen zu können. Dabei wird auf das Prinzip zurückgegriffen das innere Seelenleben, wie das eigene Denken, Fühlen und Empfinden zu beobachten und daraufhin Rückschlüsse zu ziehen. [BIS09]

Der Proband ist dabei gleichzeitig der Beobachtende sowie der Beobachter. Er selber ist es, der Auskunft geben muss darüber was empfunden/wahrgenommen wird. Und gleichzeitig ist dies auch das Problem. Die Ergebnisse dieses Versuches hängen hierbei sehr stark ab von der Fähigkeit des Probanden sich so auszudrücken wie es dem Experiment dienlich ist. Der Proband muss selber qualitativ entscheiden wann der gewünschte Zustand erreicht wurde. Eine quantitative Erfassung gewünschter Zustände ist hierbei wesentlich schwieriger.

Der deutsche Psychologe Wilhelm Wundt vollzog seine Experimente daher sehr stark an die Naturwissenschaften angelehnt. Um eine möglichere quantitative Erfassung zu ermöglichen verwendete er bei seinen Experimenten optische Verschlüsse oder Geräte die eine elektrische Zeitmessung ermöglichten. Das Experiment bestand bei ihm darin, dass äußere Reize wie optische, taktile oder akustische Einflüsse einige Sekunden auf den Probanden einwirken zu lassen, worauf dieser zu diesen Reizen hin assoziieren sollte was er „vor seinem inneren Auge sieht“. Die Frage die Wundt klären wollte war: „wie ist der assoziative Verlauf von Vorstellungen?

Der deutsche Psychologe Karl Bühler verfolgte einen ganz anderen Weg. Er versuchte mit Hilfe von Experimenten einen besseren Einblick in das Denken zu bekommen. Den Versuchspersonen wurden recht schwierige Fragen oder Sätze gegeben, sodass die Zeit bis zur Antwort etwas länger ausfallen konnte. Nach einer gestellten Frage hatten die Personen die Aufgabe ihre Empfindungen und Gedanken bis zur Antwort kundzutun.

Nach all diesen Versuchen wollte Bühler damit Aussagen über die Bestandteile und Strukturen von Prozessen des Denkens treffen und über die Konstitution dieser Bestandteile. In einem anderen Experiment ließ er Patienten mit verbundenen Augen Dinge ertasten. Der Proband sollte dabei immer seine Assoziation zu dem was er ertastet mitteilen. Bei dem Ertasten einer Haarnadel war die erste Assoziation die eines gekrümmten Drahtes. Nach kurzer Zeit sank die Vorstellung des gekrümmten Drahtes, die der Proband durch das Tasten bekam. Dafür steigerte sich aber die Gesichtsvorstellung mit der Zeit, sodass er anhand der Form des Drahtes diesem viel mehr als Haarnadel identifizieren konnte.

Ziel von Wundt war es dabei Aussagen und Erkenntnisse über den Verlauf der Entwicklung von Vorstellungen zu bekommen, sowie die Bestandteile und Strukturen von Prozessen des Denkens und zu erkennen wie der Zusammenhang von Reiz zur Wahrnehmung und Empfindung sich äußert. [BIS09]

Der Konstruktivismus ist eine Richtung der Psychologie, die das Konstruieren einer eigenen Vorstellung der Umwelt durch die wahrgenommenen Reize eben jener Umwelt zu erklären versucht. Ein klassischer Vertreter des Konstruktivismus ist Emanuel Kant, der auch eine der Hauptpersonen der Aufklärung war.

Der Konstruktivismus folgt dabei einem bestimmten Leitspruch:

*„Der Mensch hat keinen unmittelbaren Zugriff auf die objektive Realität,  
wahr ist, was (als) wahrgenommen wird“ (Robak, 2012)*

Vereinfacht gesagt wird festgelegt, dass komplexe Lebensformen die Umgebung nicht selber direkt aufnehmen, sondern vom Zeitpunkt ihrer ersten Wahrnehmung an Reize aufnehmen und diese zu einer eigenen Vorstellung der Umgebung ausbauen. Das bedeutet, dass es zwar eine Wirklichkeit gibt, sie aber nicht genau erfassbar ist, sondern für jeden Menschen, abhängig von den Erfahrungen und der bisherigen Vorstellung der Umwelt anders aussieht. Es muss also unterschieden werden zwischen Wirklichkeit und Realität. Wahrheit ist subjektiv und vom Menschen abhängig wohin die Realität aber objektiv ist. Die Realität ist das worin der Körper in der Umwelt sich aufhält. Wahrheit ist das, was das Subjekt aus den gewonnen Daten sich konstruiert um seine Umgebung zu beschreiben. Tritt der Mensch nun in eine Situation ein, für die er keine Vorstellung konstruiert hat, weil er zum Beispiel noch nie an diesem Ort war, wird er orientierungslos und unsicher und muss in der neuen Umwelt erst einmal eine neue Vorstellung konstruieren. Jegliches Verständnis des Menschen für bestimmte Sachverhalte ist demnach abhängig von dem inneren Modell des Menschen, was gleichbedeutend ist mit der konstruierten Vorstellung seiner Umgebung. Passt der Inhalt dessen was er neu lernen muss nicht in seine Vorstellung, wird er diese Sachverhalte nur schwer erlernen können, bis er sein inneres Modell auf die neuen Sachverhalte hin erweitert hat.

Der Konstruktivismus ist demnach zum Teil Behaviorismus und zum Teil Introspektion. Er beinhaltet sowohl die Aspekte der äußeren Betrachtung und der Reaktion des Subjektes, als auch die innere Betrachtung des Subjektes selber. Dabei unterscheidet man zwischen dem radikalen und dem sozialen

Konstruktivismus. Der radikale Konstruktivismus besagt, dass jedes Individuum sich seine Umgebung selber gestaltet. Der soziale Konstruktivismus besagt jedoch, dass Informationen zur Umgestaltung der eigenen Vorstellung durch andere Individuen erreicht werden kann. [BIS09]

Die Annahme, dass das Individuum die Wirklichkeit nicht erfassen kann stößt dabei auf Kritik. Denn das nicht zurückgreifen auf die Wirklichkeit, kann keine Rückschlüsse auf die Wirklichkeit ziehen lassen.

### 2.3 Biologische Perspektive

Nach den bisher ausgeführten Richtungen der Psychologie ergibt sich eine klare Trennung des äußeren und des Inneren. Wie bereits beschrieben richtet der Behaviorismus seinen Fokus auf die Reaktionen eines Systems auf äußere Einflüsse. Die Introspektion fokussiert sich dagegen jedoch auf die Innere Ansicht des Systems welches ihn zu seiner Reaktion bringt. Aus Biologischer Sicht richtet jedes System seine Reaktionen auf äußere Einflüsse immer auf einen gewissen Zweck hin aus.

Descartes selber trennt das innere und das äußere als zwei Substanzen die miteinander agieren aber getrennt voneinander sind. Das denkende Ich selber bezeichnet er als „res cogitans“, eine Substanz die mit dem körperlichen selber absolut nichts zu tun hat.

Der gesamte Körper und alle Reaktionen die in ihm stattfinden wird als „res extensa“ bezeichnet. Die Darstellung beider Ebenen getrennt zu einander kann mit Abbildung 1 veranschaulicht werden. Beide Substanzen interagieren mit einander, was als Dualismus bezeichnet wird. [BIS09]

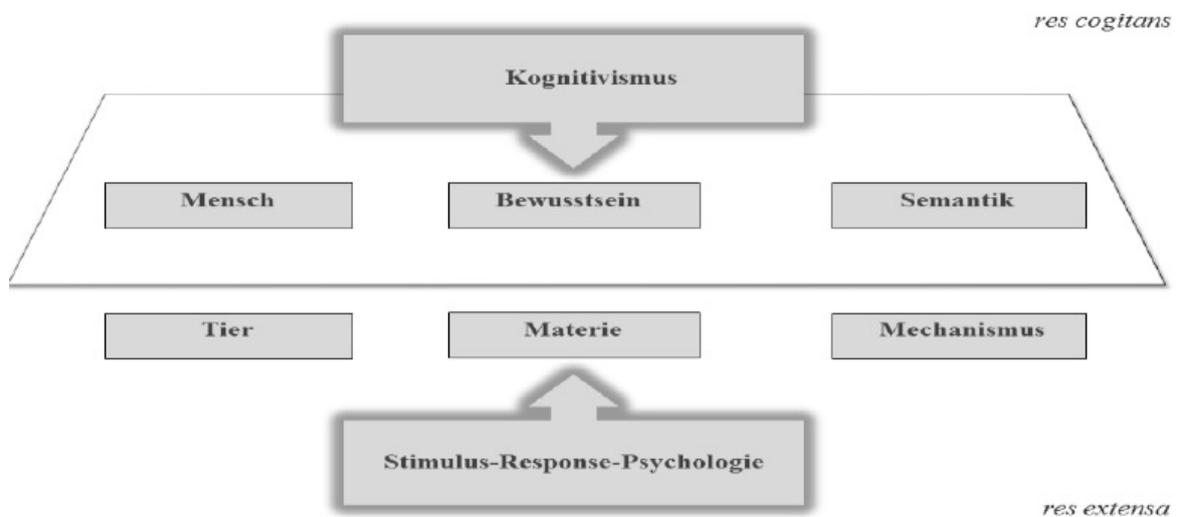


Abbildung 1: Darstellung innere gegen äußere Sicht [BIS09]

Die untere Ebene der Abbildung stellt die Aspekte dar, die im Behaviorismus betrachtet werden. Dabei wird das materielle, also der reine Körper und seine Reaktionen auf äußere Einflüsse betrachtet, die *res extensa*.

Der Kognitivismus betrachtet die Elemente der oberen Ebene der Abbildung. Hierbei wird der Mensch vom Tier abgegrenzt, da ihm höhere geistige Fähigkeiten zugesprochen werden.

Die Philosophen Leibniz und Spinoza erklären diese beiden jedoch nicht als Dualismus, sondern als Parallelismus. Das Denkende ich und das Materielle ich stellen die selbe Substanz dar, jedoch von zwei verschiedenen Seiten.

Um diesen Dualismus aufzulösen verlangt Bischof eine Betrachtung der biologischen Sicht. Diese Sicht verlangt eine Betrachtung der äußeren Umstände sowie der inneren Perspektive. Die Notwendigkeit beide Perspektiven zu betrachten ergibt sich schon aus dem verschiedenen Verhalten von Tieren bei gleichen äußeren Einflüssen. Zusätzlich zu der rein äußeren Betrachtung muss nun eine semantische Ebene folgen, in der die Organismen eingehenden Reizen eine Bedeutung zuweisen. Der Vorteil dieser Bedeutungszuweisung wird im nächsten Kapitel genauer beschrieben.

Durch die Erweiterung der semantischen Ebene ist es nun möglich eine physische und kognitive Entwicklung von Organismen zu beschreiben. Es bietet Ansätze um antrainiertes Verhalten von Kindern beim Aufwachsen zu erklären oder das Artverhalten von Organismen über die Evolution hinweg.

Berücksichtigt man diesen biologischen Aspekt, dann besteht offensichtlich kein Dualismus mehr zwischen dem Materiellem und dem Geistigen. Vielmehr kristallisiert sich eine Entwicklungspyramide wie in Abbildung 2 dargestellt heraus. Dazu wird in der untersten Ebene der rein mechanische Körper dargestellt, der rein nach dem Prinzip des Behaviorismus auf äußere Reize vorhersagbare Reaktionen zeigt. Jegliches Verhalten in dieser Ebene ist rein reflexiv zu begründen.

Die nächste Ebene weist den eingehenden Reizen eine Bedeutung zu und kann den Organismus vor einer Terminierung schützen. Diese Ebene ist für Organismen von Bedeutung, um ihr Überleben zu gewährleisten. Über diese Ebene hinaus haben sich Organismen entwickelt die über ein Bewusstsein verfügen, die Emotionen und Erfahrungen machen können und so nach der Introspektion ihre eigene phänomenale Welt aufbauen können. Der Mensch selber hat sich zusätzlich noch dazu die Fähigkeit entwickelt Erkenntnisse zu schaffen und seine Umgebung zu beeinflussen und gewonnene Erkenntnisse zu nutzen um seine Umgebung zu manipulieren.

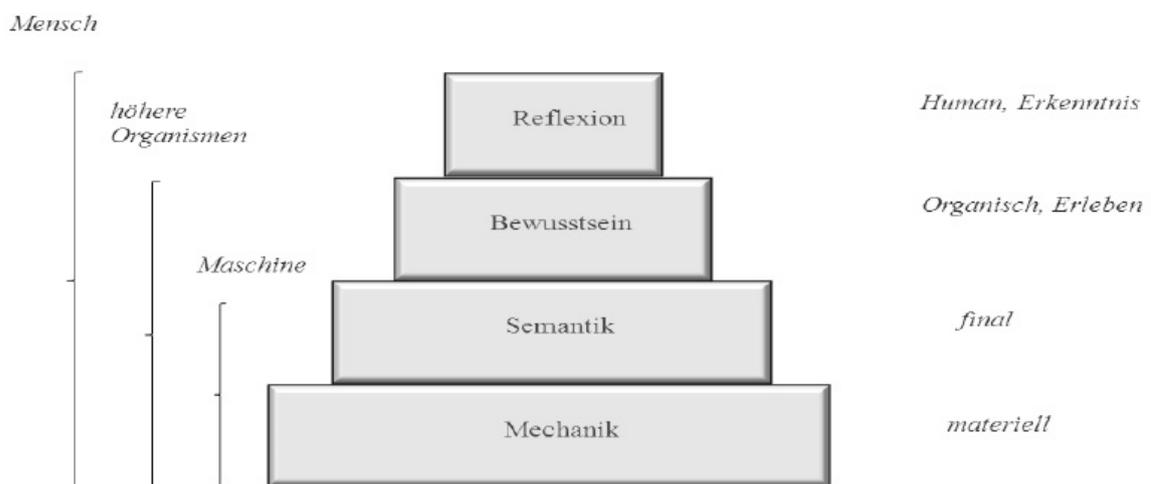


Abbildung 2: Entwicklungspyramide [BIS09]

## 2.4 Finale Systeme

Kein reales System ist komplett von seiner Umwelt isoliert, es steht immer im Kontext zu seiner Umgebung. Das System nimmt ein Ereignis in der Umgebung über ein physikalisches Signal der Umgebung auf. Jedoch hängt es von System ab, welche Bedeutung diesem Signal beigemessen wird. Ein Signal, das vom System aufgenommen wird, kann seine Bedeutung mit der Zeit ändern. Dies wird in der Psychologie Anpassung oder Adaption genannt. Gewisse Verhaltensmuster sind von Geburt an vorhanden, wie z.B. etwaige Schutzreflexe, die das Subjekt davor schützen soll Gefahren zu vermeiden. Junge Tiere verhalten sich gegenüber einer unbekanntem Rasse, da sie die fremde Art nicht einschätzen können. Junge Katzen haben vor Menschen Angst, da sie keine Erfahrung mit diesen aufzuweisen haben. Wird das Tier aber gut behandelt, also macht positive Erfahrungen mit dem Menschen, wird es, wenn es einen Menschen sieht, die Bedeutung dieses Signals mit der Zeit adaptieren und sich Zutrauen zu der anderen Art entwickeln.

Ähnlich wie bei LTI-Systemen in der Technik, bei denen man das System einem Signal oder Reiz aussetzt und die Reaktion auf das System untersucht um auf die Zusammenhänge des Systems zu schließen, können auch reale Systeme auf die Art geprüft werden und kausale Zusammenhänge formal beschrieben werden. Diese Funktion, ähnlich wie im vorherigen Kapitel beschrieben, stellt Zusammenhänge dar von dem Eingangereignis A zur Ausgangsreaktion B. Jedoch gibt es einige Reaktionen, die die Existenz des Systems beenden würden. Um dies zu verhindern existiert keine Funktion die als Folge eines solchen Reizes zu solchen finalen Reaktionen führt. Dies würde bedeuten, dass das System „sterben“ könnte also seine Existenz begrenzt ist. Daher wurde der Ausdruck finales System eingeführt. Ziel des Systems ist es sein Verhalten vom Reiz zur Reaktion hin zu optimieren, was als Adaption bezeichnet wird.

Das System befindet sich stets in Wechselwirkung mit seiner Umgebung. Es empfängt von dieser Umgebung dauerhaft physikalische Signale, welche die Umgebung beschreiben und es selber reagiert in dieser Umgebung und

beeinflusst sie durch die Aktionen, die das System vollzieht mit. Demzufolge besitzt jedes einzelne Signal, welches von der Umgebung gesendet wird, eine eigene Bedeutung, welche als Designat bezeichnet wird [BIS09]. Jedes System verfügt über andere Grenzen, die seine Finalität bestimmen. Daraus folgt, dass jedes System, die Signale die es bekommt anders deuten oder interpretieren muss. Damit wird die rein mechanische Ebene aus Abbildung 2 erweitert um eine semantische Ebene. Die Semantik ist also für finale Systeme von essentieller Bedeutung, da sie Signale aus der Umgebung nicht nur aufnehmen, sondern auch bewerten müssen um zu verhindern, dass sie sich in einen Zustand begeben, der sie schaden könnte. In der Natur weist jeder Organismus einem Reiz eine Bedeutung zu. Diese Semantik bestimmt das Verhalten des Organismus. Ein wildes Tier würde bei dem Auftauchen eines Menschen eine Gefahr implizieren und die Flucht ergreifen. Bei einem domestizierten Tier ist das Erscheinen eines Menschen nicht unbedingt negativ, weswegen das Tier nicht automatisch wegrennt.

Bischof beschreibt diesen Sachverhalt mit dem Verhalten von Affen die einer Schlange begegnen. Durch das Aussehen der Schlange erlangt der Affe die Informationen über die Giftigkeit des Tieres. Der Affe deutet also auf semantischer Ebene die Informationen, die er visuell wahrnimmt. Abhängig von der Bedeutung, die er der Schlange zuweist, wird er daraufhin sein Verhalten festlegen. Deutet er die Schlange als giftig, wird der Affe schnell die Flucht ergreifen. Welche Bedeutung der Affe dem Aussehen der Schlange zuweist, hängt dabei von seinen Erfahrungen sowie dem bereits erlangten Wissen ab.

Nach der Grundlage des Konstruktivismus würden Organismen, die über eine semantische Ebene verfügen, sich aus ihren Erfahrungen eine eigene Vorstellung der Welt konstruieren. Das ist für finale Systeme ein essentiell wichtiges Element, da mit einem inneren Modell der Welt Zustände vorhersagbar sind, die einen gefährden könnten. Somit können Organismen verhindern in gefährliche Zustände zu kommen, da sie diese schon in ihrer inneren Vorstellung bewerten können bevor sie diese in der Realität wirklich umsetzen.

### 3. Technische Modellierung kognitiver Systeme

#### 3.1 Der kognitive Kreis

Da die psychologischen Betrachtungen auch auf technische Systeme jetzt angewendet werden sollen, wird im Folgenden nicht mehr zwischen technischen System oder Organismen unterschieden, sondern nur noch von Agenten gesprochen. Deswegen wird nun der Begriff des Agenten einmal genauer erklärt.

Ein Agent ist alles, was seine Umgebung durch Sensoren wahrnehmen kann und in dieser Umgebung durch Aktuatoren handelt [RUS04]. Menschen und Tiere verfügen über nahezu gleiche Sensoren, wie die Augen zum Sehen oder die Ohren zum Hören, jedoch sind diese abhängig von ihrem Lebensraum verschieden ausgeprägt. Um Aktionen auszuführen gibt es verschiedene Aktuatoren, wie z.B. die Gliedmaßen zum Bewegen.

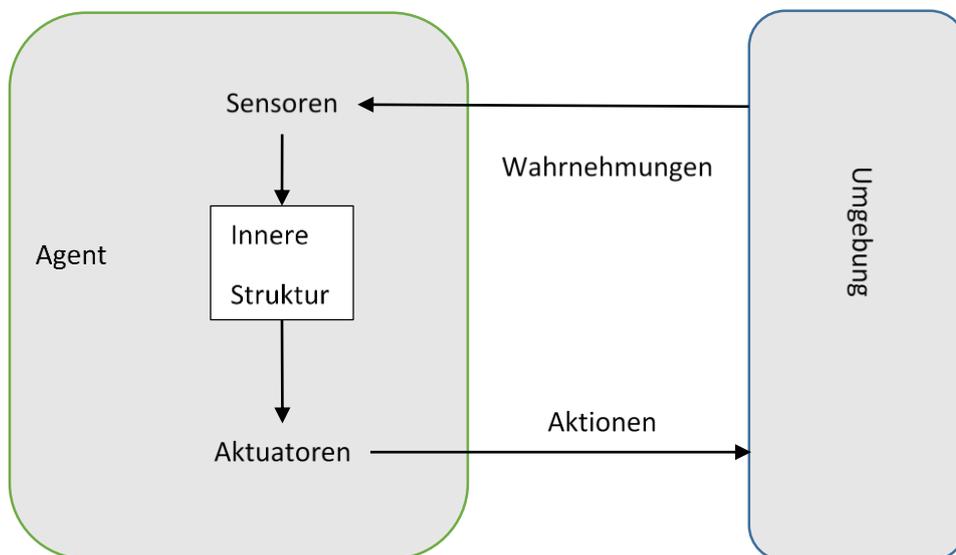


Abbildung 3: Relation zwischen Umgebung und Agenten

Technische Systeme besitzen mehrere Zustände in denen sie sich aktuell befinden und Übergänge, welche die Bedingungen angeben unter denen das System von einem Zustand in einen anderen wechselt. Aufgrund dessen ist es zielführend kognitive Abläufe zu formalisieren.

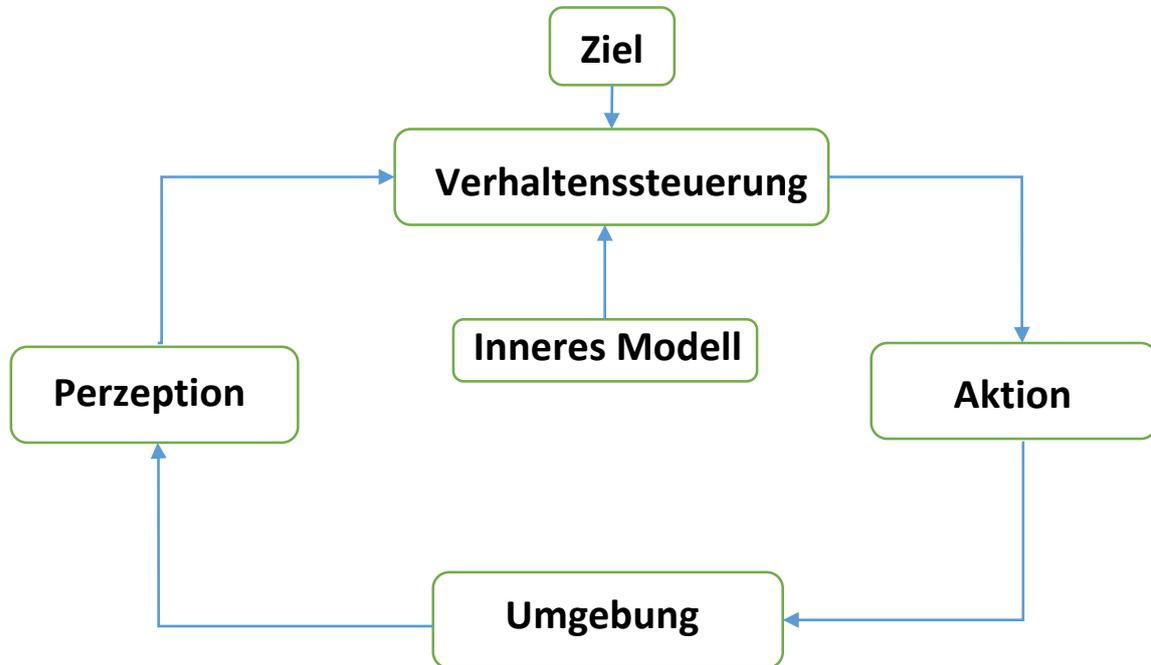
Betrachtet man nun die Prinzipien eines finalen Systems genauer spielen vier Aspekte für den Agenten eine Rolle.

Die Umgebung wird vom Agenten durch seine Sensoren wahrgenommen. Verschiedene Organismen nehmen abhängig von ihrem Lebensraum und ihren Beuteschemas gewisse physikalische Einflüsse stärker wahr als andere. Die Ohren der Fledermaus nehmen einen viel höheren Frequenzbereich wahr als die des Menschen, da Fledermäuse ihre Ohren als Echolote verwenden. Ein Agent muss die Signale, die er aufnimmt deuten können um seine Finalität zu sichern. Diese Signaldeutung wird als Perzeption bezeichnet. Nach dem der Agent seine Umgebung semantisch bewertet hat, ändert der Agent seine Aktionen um in der Umgebung, die er gerade um sich hat, sein Ziel erreichen zu können. Dabei wählt er die Strategie aus, die den größten Erfolg verspricht.

Hat der Agent nun eine Strategie gefunden ändert er sein Verhalten und muss nun eine passende Aktion auswählen um seine Strategie in der Umgebung umsetzen zu können. Seine Handlungen aber bewirken nach der newtonschen Mechanik, also nach dem dritten Newtonschen Axiom „Accio = Reaccio“, immer eine Konsequenz. Jede Aktion die ausgeführt wird, beeinflusst die Umgebung wieder, sodass der Agent die veränderte Umgebung neu wahrnehmen muss.

Der Elektroingenieur Simon Haykin entwickelte das Modell eines kognitiven Kreises, wie in Abbildung 4 dargestellt, in welchem er den kognitiven Prozess der finalen Systeme als Kreis darstellt mit den eben beschriebenen Komponenten. Der Kreis umfasst auf der unteren Hälfte der Abbildung die rein mechanische Ebene in der der Agent durch seine Aktuatoren in der Umgebung wirkt und durch seine Sensoren er die Signale der Umgebung auffasst. Die obere Hälfte gehört zur rein semantischen Ebene, in der er die Signale, die er von der Umgebung bekommt auffasst und in ein Verhalten, abhängig von seiner Weltvorstellung und dem Ziel, das er erreichen will, überführt.

## Semantische Datenstrukturen



## Signalebene

Abbildung 4: Kognition nach Heykins

Dieser kognitive Kreis ist eine strukturierte Darstellung dessen, was ein finales System voraussetzt. Das Verhalten eines Organismus wird immer durch die Umgebung mitbestimmt und besagt aber auch, dass das Verhalten eines Organismus nach äußeren Reizen begrenzt vorhersagbar ist.

Um das Ganze zu veranschaulichen, wird dieser Kreis am Beispiel eines Frosches verdeutlicht.

Angenommen, ein Frosch befindet sich im Wald, wo er im dichten Grün auf seine Beute wartet. Dort sitzend wird er regungslos verharren, da sich keine Beute bisher genähert hat. Das bedeutet, der Agent, in diesem Fall der Frosch, wird keine andere Aktion durchführen, solange sich die Umgebung nicht ändert. Bei gleichbleibender Umgebung wird der Frosch immer dasselbe wahrnehmen, womit

er immer dasselbe Verhalten an den Tag legen wird.

Erst wenn sich die Umgebung ändert, ändert sich auch die Wahrnehmung des Tieres und damit auch sein Verhalten. Nähert sich zum Beispiel eine Fliege wird der Frosch sein Verhalten so ändern, dass er die Fliege als Beute verspeisen kann. Nähert sich zum Beispiel ein Storch, wird der Frosch sein Verhalten von Beute jagen ändern dahin, dass er sich selbst retten will und wird möglichst das Weite suchen wird um nicht selbst zur Beute zu werden. Das dieses Beispiel nicht vollkommen der beobachteten Realität entspricht, wird an dieser Stelle schnell klar. Im Kapitel vier wird dazu näher eingegangen um die Probleme des einfachen kognitiven Kreises darzustellen.

## 3.2 Lernformen

### 3.2.1 Biologisch orientiertes Lernen

Bei der Beschreibung des Lernprozesses wird sich stark an dem Prinzip von finalen Systemen angelehnt. Der Agent besitzt eine endliche Anzahl an Zuständen die er erreichen kann. Sind die Zustände die der Agent erreicht näher an dem Ziel wird dem Agenten eine „Belohnung“ gegeben werden, z.B. ein hoher Zahlenwert, um dieses Verhalten, also den Wechsel in diesen Zustand positiv zu bewerten. Entfernt er sich von dem Ziel bekommt er eine „Strafe“, also einen niedrigen Zahlenwert um sein Übergang in den Zustand negativ zu bewerten. Der Agent kann seine Zustände erforschen in dem er sich in einen zufälligen Zustand begibt und eine Aktion auswählt. Ist diese Aktion zielführend wird diese positiv verstärkt in dem sie belohnt wird. Dies wird als Reinforcement Learning bezeichnet.

Dies kann zum Beispiel an dem von Thorndike postuliertem psychologischen Verfahren des „operanten Konditionieren“ verdeutlicht werden.

Tiere oder auch Menschen, werden auf die Art trainiert, dass negative Verhaltensweisen bestraft werden und gewünschte Verhaltensweisen belohnt werden. Im Allgemeinen wird diese Form der Sozialisierung auch als Erziehung bezeichnet.

Dieses Prinzip ist ein wichtiger Bestandteil von finalen Systemen, da hierbei durch das Zuführen bestimmter Reize auch die innere Vorstellung des Subjektes mit beeinflusst wird und so das Verhalten des Subjektes mit beeinflusst wird.

Der Agent verfolgt immer ein bestimmtes Ziel, das er in seiner Umgebung erreichen möchte. Da er eine eigene Vorstellung der Welt besitzt, vollzieht er eine Aktion, die seinem Modell nach am ehestem zum Ziel führen wird. Seine Aktion in der Umgebung bewirkt eine Reaktion der Umgebung. Diese Reaktion bewertet der Agent ob diese für die Erreichung seines Ziels positiv oder negativ ist. Diese Deutung der Reaktion ist gleichzusetzen mit einer Belohnung, die der Agent von der Umgebung bekommt.

Somit kann der Agent die Auswirkung seiner Aktion deuten und wird so seine eigene Vorstellung der Welt anpassen um das Erreichen seiner Ziele zu optimieren.

Demnach gleicht das Subjekt dauerhaft seine Vorstellung der Welt mit den aus den Reizen gewonnenen Bedeutungen ab. Stimmen diese beiden überein wird der Agent immer dasselbe Verhalten ausüben. Stimmen diese nicht überein, muss das Subjekt ein neues Verhalten entwickeln um sein Ziel optimal zu erreichen.

Dieser Vergleich der inneren und äußeren Welt wird in der Psychologie unter dem Begriff der Veridikalität geführt. Veridikalität bedeutet sinngemäß so viel wie „durch Erfahrung bestätigt“. In der Praxis bedeutet dies, dass der Organismus Erfahrungen gemacht hat was seine eigene Vorstellung der Welt prägt. Tritt der Organismus nun in einen Zustand ein den er selber noch nicht kennt, sind die innere und die äußere Welt des Organismus nicht mehr veridikal. Dies hat zur Folge, dass der Organismus seine innere Welt anpassen muss, damit neu gewonnene Erfahrungen sich in diese Welt eingliedern können.

### 3.2.2 Formalisierung des Reinforcement Learning

Der Agent hat ein vorgegebenes Repertoire an Aktionen  $a$  die er ausführen kann. Da der Agent ein bestimmtes Ziel verfolgt, wird das Ausführen jeder Aktion mit einer Belohnung versehen, die das näher kommen an das Ziel indiziert. Führt er eine Aktion aus bekommt er von der Umgebung seinen neuen Zustand  $z$  aus der Übergangsfunktion  $\delta$  und eine Belohnung  $r$  von der Umgebung wieder.

$$z' = \delta(z_t, a_t) \quad (3.1)$$

$$r_t = r(z_t, a_t) \quad (3.2)$$

Damit der Agent zu seinem Ziel kommt benötigt er eine Strategie  $\pi$  die sich aus den Belohnungswerten  $r(z_t, a_t)$  ergibt. Die Optimale Strategie  $\pi^*$  ergibt sich aus den maximalen Belohnungswerten. [ERT09]

Demzufolge kann man jeden Zustand mit einer abgeschwächten Belohnung versehen, sodass sich eine Belohnungsmatrix gibt, die die selbe Dimension hat wie die Zustandsmatrix selber, da jeder Zustand einen Wert als Belohnung bekommt. Die Belohnung, die ein Zustand, bekommt richtet sich nach der direkten Belohnung, also die Belohnung, wenn der Zustand sich um den Zielzustand herum befindet, und den maximalen Belohnungen der Folgezustände. Die maximale Belohnung des Folgezustands wird jedoch in abgeschwächter Form, durch einen Diskontinuitätsfaktor  $\gamma$ , in die eigene Belohnung eingehen, damit eine Unterscheidung der Belohnungen vom aktuellen Zustand zum anderen möglich ist. Diese Art des Lernens wird als V-Lernen bezeichnet.

So ergibt sich die Belohnungsverteilung zu:

$$V(z_t) = r_t + \gamma \cdot r_{t+1} + \gamma^2 \cdot r_{t+2} + \dots = \sum_{i=0}^{\infty} \gamma^i \cdot r_{t+i} \quad (3.3)$$

Jedem Zustand werden Belohnungen zugewiesen abhängig von den umliegenden Folgezuständen. Da aber nicht alle Folgezustände zielführend sind, werden nur die Aktionen ausgewählt, die in die Zustände mit maximaler Belohnung führen würde, da diese am nächsten zum Ziel führen würden.

$$V^*(z_t) = \max_{a_t, a_{t+1}, \dots} (r_t + \gamma \cdot r_{t+1} + \gamma^2 \cdot r_{t+2} + \dots) \quad (3.4)$$

Diese Maximierung kann auch auf die einzelnen Teile der Funktion angewendet werden.

$$V^*(z_t) = \max_{a_t} [r_t + \gamma \cdot \max_{a_{t+1}, a_{t+2}, \dots} (r_{t+1} + \gamma^2 \cdot r_{t+2} + \dots)] \quad (3.5)$$

Da die Maximierung in der Klammer der optimalen Folgebelohnung entspricht kann folgende rekursive Funktion formuliert werden

$$V^*(z_t) = \max_{a_t} [r_t + \gamma \cdot V^*(z_{t+1})] \quad (3.6)$$

Da der neue Zustand  $z_{t+1}$  durch die Übergangsfunktion beschrieben werden kann folgt daraus:

$$V^*(z) = \max_a [r_t + \gamma \cdot V^*(\delta(z, a))] \quad (3.7)$$

Dies bedeutet, dass jeder Zustand in dem sich der Agent befinden kann eine Belohnung zugewiesen bekommt, die sich aus der direkten Belohnung von dem eigenen Zustand ergibt zuzüglich der verminderten maximalen Belohnung aus den erreichbaren Folgezuständen in die er übergehen würde.

Diese erlernte V-Matrix beschreibt die innere Vorstellung des Agenten, die er von der Welt erlangt hat. Wenn er nun sein Ziel erreichen möchte muss er nun in die Zustände übergehen, die ihm die maximale Belohnung bieten würden. Für die optimale Strategie werden aus der V-Funktion die Aktionen gewählt, die in die Zustände mit maximalen Belohnungen führen würden.

$$\pi^*(z) = \arg \max_a [r(z, a) + \gamma \cdot V^*(\delta(z, a))] \quad (3.8)$$

Der gesamte Zustandsraum wird so oft durchlaufen bis die gesetzten Belohnungswerte für die V-Matrix konvergieren und somit ist dann die optimale V-Matrix entstanden

Der entscheidende Punkt beim V-Lernen ist, dass jeder Zustand eine Belohnung bekommt nach der sich orientiert wird beim dem Entwickeln der Strategie.

Deswegen wird nun vom V-Lernen in das Q-Lernen übergegangen.

Beim Q-Lernen wird sich nicht nur nach der Belohnung des Nachfolgezustandes orientiert, sondern es werden Belohnungen für jede Aktion vom aktuellen Zustand verteilt. Nach dem Maximum der Belohnung wird dann nicht der Nachfolgezustand bestimmt, sondern die nächste Aktion. Dies hat den Vorteil, dass nach dem Bestimmen einer Aktion der Folgezustand noch nicht betreten wurde und überprüft werden kann, ob der Zustand sicher ist zum Betreten.

Das Lernen selber findet äquivalent zum V-Lernen statt, wobei nun hier nicht der Zustand eine Belohnung bekommt, sondern die mögliche Aktion des Zustandes.

$$Q(z_t, a_t) = \max_{a_t, a_{t+1}, \dots} (r(z_t, a_t) + \gamma \cdot r(z_{t+1}, a_{t+1}) + \gamma^2 \cdot r(z_{t+2}, a_{t+2}) + \dots) \quad (3.9)$$

Die Maximierung der Werte kann nun wieder auf die Folgebelohnungen erweitert werden

$$Q(z_t, a_t) = r(z_t, a_t) + \gamma \cdot \max_{a_{t+1}, a_{t+2}, \dots} (r(z_{t+1}, a_{t+1}) + \gamma \cdot r(z_{t+2}, a_{t+2}) + \dots) \quad (3.10)$$

Die zweite Maximierung in der Formel ist gleichbedeutend mit der vorherigen Q-Matrix im Zustandsraum

$$Q(z, a) = r(z, a) + \gamma \cdot \max_{a'} Q(\delta(z, a), a') \quad (3.11)$$

Praktisch funktioniert dies so, dass der Agent einen zufälligen Zustand betritt und eine zufällige Aktion auswählt und für diese Aktion die maximale Belohnung aus dem Folgezustand dieser Aktion auswählt. Somit orientiert sich der Agent am Ende seines Lernprozesses an der erlernten Q-Matrix aus denen er sich die notwendigen Aktionen ableitet anstatt die nächsten Zustände anzugeben.

$$\pi(z) = \mathit{arg} \max_a Q(z, a) \quad (3.12)$$

### 3.2.3 Adaption und Coping

Adaption vom englischen Wort „to adapt“ bedeutet sich anpassen. Auf die Thematik bezogen heißt dies, dass der Agent sich ebenso an Veränderungen in seinem Zustandsraum anpassen lernen muss.

Der Agent hat seine Umgebung gelernt in dem er Belohnungen für jede Aktion in jedem Zustand verteilt hat. Da die Umgebung sich ändern kann, kann es passieren, dass für eine Aktion in einem bestimmten Zustand die Übergangsfunktion nicht mehr dasselbe Ergebnis liefert wie vorher.

Da der Agent sich aber bei der Wahl seiner Aktionen sich nach den maximalen Belohnungen richtet, müssen diese Belohnungen angepasst werden. Dabei muss der Agent die Umgebung neu lernen um die Belohnungen neu zu verteilen.

Da aber nur die Zustände nahe des Hindernisses eine hohe Abweichung in den Belohnungswerten aufweisen, sind beim neuen Lernen des Zustandsraumes viel weniger Schritte erforderlich als beim Lernen des kompletten Zustandsraumes. Dieses Verfahren wird schon automatisiert von uns im Alltag angewendet. Laufen wir beispielweise unsere gewohnte Strecke von A nach B und befindet sich

unerwartet ein Hindernis, z.B. eine gesperrte Straße, gehen wir den Weg um das Hindernis herum, welchen den meisten Erfolg verspricht. Befinden wir uns dabei in einer fremden Stadt, würden wir uns eine Karte nehmen und die Umgebung neu erlernen.

Coping ist ein Vorgang aus der Psychologie, der auch unter dem Begriff „Bewältigungsstrategie“ geführt wird. Hierbei geht es darum ein Hindernis, was ein Individuum im Weg liegt zu bewältigen mithilfe der von ihm bereits bekannten Aktionen.

Beim Coping werden die bekannten Aktionen zu neuen Aktionen kombiniert um das Hindernis zu bewältigen. Stellt man sich einen Roboterarm vor, der seine Bewegungen in nur zwei Richtungen also horizontal und vertikal durchführen kann, wäre eine neue Aktion eine diagonale Bewegung, also in beide Richtungen. Da diese Aktion unbekannt ist, liegen keine Belohnungen für diese Aktionen vor, sodass der Zustandsraum noch einmal gelernt werden muss, jedoch nur für diese neu gelernte Aktion.

Da ein Hindernis etwas Unbekanntes darstellt, ist dem Agent auch nicht bekannt welche Kombination seiner Aktionen zielführend ist. Demnach kombiniert er seine bekannten Aktionen und vollzieht Probehandlungen um zu prüfen ob er mit dieser das Hindernis überwinden kann.

An dem Beispiel mit der Straßenblockade würde dies bedeuten, dass wir abschätzen ob man durch eine geeignete Anordnung aller bekannten Aktionen einen Weg durch die Blockade finden können. Möglich wäre es eventuell über die Blockade zu klettern oder sich durch freie Lücken durch zu bewegen.

Um die neue Aktion zu ermitteln betrachtet der Agent sowohl seine Aktion im aktuellen Zustand, als auch die Aktionen im Folgezustand. Der Agent „merkt“ sich seine Aktion und betrachtet nun die Aktion mit der maximalen Belohnung im Folgezustand.

$$a' = \mathop{\text{arg max}}_a Q(z', a) \quad (3.13)$$

Beide Aktionen werden nun addiert und so die neue Aktion ermittelt.

$$a_{neu} = a + a' \quad (3.14)$$

Beinhaltet die neue Aktion nur Elemente die auch in der möglichen Aktionsmatrix enthalten sind dann ist die Aktion auch gültig. Falls nicht, muss die nächste Aktion mit der maximalen Belohnung aus dem Folgezustand genommen werden.

```

repeat
  a' = arg max_a Q(z', a)
  a_neu = a + a'
  If (a_x ∉ A) ∧ (a_y ∉ A)
    max_a' Q(z', a') = 0
  else
    A = A ∪ a'
    break
until a' = {}

```

Bei diesem Algorithmus setzt der Agent in seiner inneren Welt zwischenzeitlich die maximale Belohnung im Zustand  $z'$  für die Aktion  $a'$  auf null, damit er die Aktion mit der zweitgrößten Belohnung in der Schleife auswählen kann-

Die neue Aktion ist nun Bestandteil der Gesamtaktionsmatrix und somit kann der Agent nun den Zustandsraum mit der neuen Aktion durchlaufen.

Da der Zustandsraum schon vorher gelernt wurde, muss die Q-matrix für die neue Aktion nicht nach dem Algorithmus des Q-Lernens erlernt werden, sondern nur die maximale Belohnung aus dem Zustand entnommen werden in die die neue Aktion führt.

$$Q(z, a_{neu}) = \max_{a_t} Q(\delta(z, a_{neu}), a_t) \quad (3.15)$$

Diese Variante hat den Vorteil, dass nicht nach zufälligem Auswählen von neuen Zuständen neu gelernt werden muss, sondern die bereits gelernten Belohnungswerte auf die neue Aktion verteilt werden. Dies ist jedoch nur möglich, da die neu gelernte

Aktion sich aus den anderen bereits bestehenden Aktionen ergibt. Die benötigte Anzahl an Lernschritten ist nun gleich der Anzahl der erreichbaren Zustände.

### 3.3 Leistungsfähigkeit

Die Leistungsfähigkeit eines Agenten ist ein Maß dafür wie erfolgreich der Agent seine Aufgabe erfüllt hat. Es ist notwendig ein objektives Maß zu finden, was zum Beispiel durch eine quantitative Bewertung des Agenten erfolgen kann. Diese Leistungsbewertung hat der Agent immer zu maximieren.

Um ein quantitatives Maß zu finden bietet es sich an eine Messung des Lernverfahrens durchzuführen. Um eine repräsentative Aussage zu bekommen, muss eine Berechnungsvorschrift erstellt werden, die eine Aussage über den Lernfortschritt beschreibt. Wird der Prozess des Q-Lernens noch einmal betrachtet bietet es sich an, eine Verteilung der neu gesetzten Belohnungen darzustellen.

Beim Q-Lernen werden zufällig Zustände ausgewählt und ebenso zufällig auch Aktionen. Die Belohnung die für die gewählte Aktion vergeben wird, ergibt sich aus der abgeschwächten maximalen Belohnung im Folgezustand.

Wird das über den gesamten Zustandsraum durchgeführt sind bis auf das Ziel alle Belohnungen erst einmal null. Nach dem die Zustände um das Ziel herum besucht wurden, sind die ersten Belohnungen verteilt worden.

Ein gutes Maß wie erfolgreich das Labyrinth gelernt wurde kann hier sein, wie sehr die neu gesetzte Belohnung vom alten Wert abweicht. Dieser Wert  $d$  wird von nun an Belohnungsdifferenz genannt.

$$d = Q_{neu}(z_t, a_t) - Q_{alt}(z_t, a_t) \quad (3.16)$$

Die Belohnungswerte verteilen sich erstmals, sobald ein Zustand um das Ziel herum betreten wurde. Da die Wahrscheinlichkeit von Anfang an sich um das Ziel zufällig zu positionieren in den meisten Fällen sehr gering ist, werden die ersten

Belohnungswerte und damit auch die Differenzen bei null liegen. Da es tausende Schritte brauchen kann bis der Zustandsraum gelernt wurde ergeben sich viele Belohnungsdifferenzen. Deswegen ist es sinnvoll eine mittlere Belohnungsdifferenz über eine gewisse Anzahl von Schritten darzustellen.

$$\bar{d} = \frac{1}{n} \sum_{t=0}^n Q_{neu}(z_t, a_t) - Q_{alt}(z_t, a_t) \quad (3.17)$$

Sobald die ersten Belohnungen verteilt werden, sind die mittleren Belohnungsdifferenzen sehr hoch. Je mehr diese sich verteilen konvergieren diese Differenzen allmählich gegen null. Diese Lernkurve wurde in Abbildung 5 dargestellt. Wie man sieht verteilen sich die meisten Belohnungen nicht von Anfang an, da zufällig die ersten Zustände gewählt werden. Je mehr sich die Belohnungen verteilen werden die Abweichungen auch geringer, weswegen diese Kurve als Maß für die Leistung des Lernens genommen werden kann.

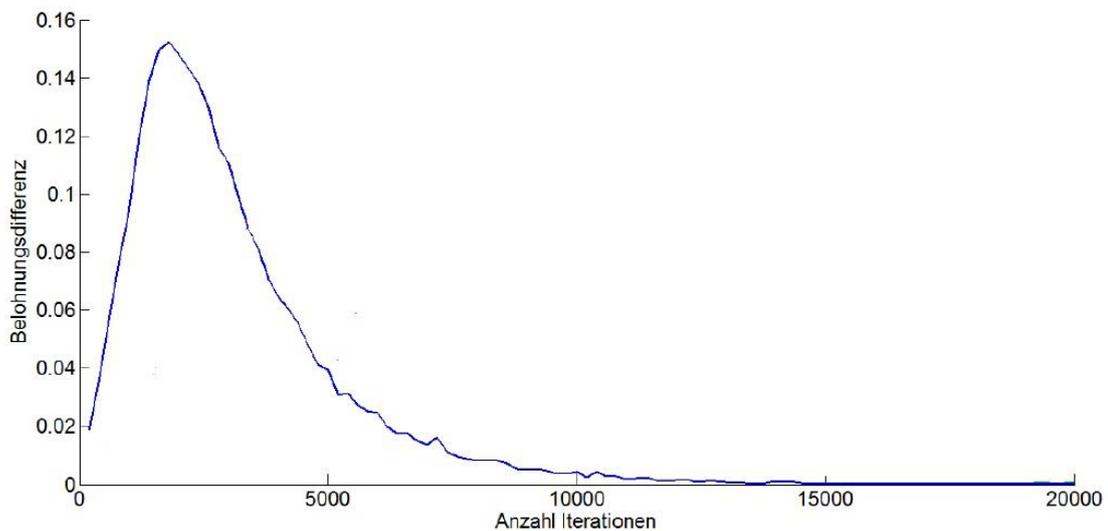


Abbildung 5: Verlauf der Lernkurve des Agenten [MEY15]

Am Beispiel der Abbildung 5 kann gezeigt werden, dass 15.000 Schritte benötigt würden bis die Funktion annähernd gegen null konvergiert ist. Der Abfall dieser Leistungskurve kann als Maß für den Lernerfolg genommen werden.

## 4. Kritik am traditionellem kognitivem Kreis

Der kognitive Kreis dient für einfache Systeme als gute Möglichkeit Kognition auf Basis des Behaviorismus zu beschreiben. Jedoch fehlt ein wichtiger Aspekt der Kognition, was man allgemein in Richtung Bewusstsein werten kann.

Wie im Kapitel zwei über den Konstruktivismus bereits beschrieben wurde, bauen sich alle kognitiv höher entwickelten Lebewesen aus ihren Erfahrungen die sie gemacht haben eine Vorstellung über die Welt auf. Die Art wie diese Welt aufgebaut ist definiert die Art und Handlungen des Individuums.

Der kognitive Kreis nach Heykins stellt die biologische Perspektive in keiner Weise ausreichend dar. Es wird hierbei nicht dargestellt, dass der Organismus eigene Erfahrungen und auch eigene Bedürfnisse hat.

Es wird hierbei nicht berücksichtigt, dass der Frosch aufgrund von Hungergefühlen und den Erfahrungen die er mit der Umgebung gemacht hat, seinen Standort ändern kann um erfolgreich mehr Beute jagen zu können. Ebenso stellt der kognitive Kreis nicht da, dass der Frosch die Änderung seiner Umgebung interpretieren muss. Wenn er nicht zwischen Beute und Feind interpretieren kann ist die Überlebensfähigkeit des Frosches sehr gering.

Ein System welches alleinig auf einen äußeren Reiz reagiert und darauf eine Handlung provoziert würde sich die vorangegangenen Aktionen nur begrenzt merken können, da nach dem kognitiven Kreis Erfahrungen und gesetzte Ziele nicht berücksichtigt werden. Das System würde die daraus erfolgten Reaktionen nicht langfristig merken können. Dies würde bedeuten, dass Fehler die schon einmal getätigt wurden weiterhin möglich sind. Dies kann man bei Fliegen am besten beobachten. Das Gedächtnis einer Fliege hält lediglich 0,3 Sekunden. Wenn eine Fliege in der Wohnung nun den Reiz des Tageslichtes ausgesetzt ist und nach draußen möchte wird sie dem Instinkt dahin zu fliegen folgen. Befindet sich dort nun eine Fensterscheibe wird sie unabdingbar gegen die Scheibe fliegen. Die gewonnene Erkenntnis, dass dort ein Hindernis ist, an dem sie nicht vorbeikommt, verliert sie nach 0,3 Sekunden wieder und fliegt wieder gegen die

Scheibe. Aufgrund dieses geringen Erfahrungsspeichers, haben Fliegen eine geringere Überlebenschance, da ebenso Gefahrenquellen wieder vergessen werden.

Bei höheren Organismen muss der kognitive Agent also erweitert werden um eine Perspektive, die den inneren Bedürfnissen des Organismus entspricht und den Erfahrungen der er bereits gemacht hat. Um die Finalität einschätzen und bewerten zu können ist es notwendig, dass der Organismus oder das System, die Reize der Umwelt, die auf ihn wirken deuten bzw. interpretieren und analysieren kann. Die Ergebnisse dessen müssen das Bild was das System von der Umgebung hat, also seine phänomenale Welt, beeinflussen und ebenso muss die Vorstellung von der Welt helfen die Ergebnisse zu analysieren. Dieser Zusammenhang wird in Abbildung 6 ebenso aufgezeigt, als feinere Struktur des kognitiven Apparates.

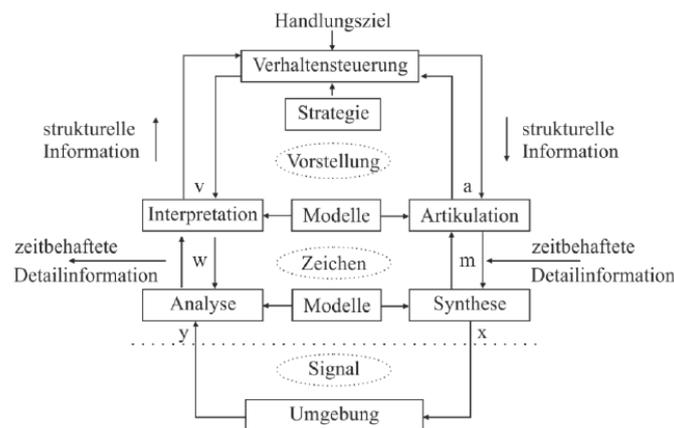


Abbildung 6: Feinstruktur eines kognitiven Agenten [RÖM13]

Die Verhaltenssteuerung muss erweitert werden um ein Ziel, welches durch das Verhalten erreichen soll, und um eine Strategie, welche helfen soll das Ziel zu erreichen. Diese Strategie, richtet sich nach der Vorstellung, die das System von der Umgebung hat. Logischerweise wird das System keine Strategie in Betracht ziehen von der es weiß, dass es in dieser Umgebung zu seinem Tod führen kann. Dies ist notwendig da höhere Organismen variable Ziele verfolgen können und

lernfähig darin sind ihre Ziele zu erreichen. Eine Maus im Labyrinth beispielsweise verfolgt das Ziel den versteckten Käse zu finden, läuft jedoch eine Katze durch das Labyrinth muss sie ihre Strategie variieren, da sie selbst ja final ist.

Hat sich der Organismus eine Strategie erstellt welches seine Finalität optimiert, muss er diese Strategie in irgendeiner Weise umsetzen. Dazu verfügt der Organismus über Fähigkeiten, die er entweder genetisch von Anfang an mitbekommen hat oder sich während seiner Lebenszeit angeeignet hat. Dies können Sprache oder spezielle Bewegungen sein die seiner Strategie zum Erfolg verhelfen sollen. Dies ist in der Abbildung 6 dargestellt durch den Bereich „Zeichen“. Als Zeichen kann jegliches semiotisches Phänomen verstanden werden, dass in irgendeiner Form dargestellt werden kann und ebenso eine spezielle Bedeutung hat. In der Tierwelt sind bestimmte Zeichen antrainiert oder erlernt. Katzen fauchen bei Gefahr oder Angst ihre vermeintlichen Gegner an um ihnen ebenso Angst zu machen und die Gegner zu verjagen. Die Form der Darstellung der erlernten Zeichen wird hierbei als Artikulation verstanden.

Da aber für den gesamten Kontext die Darstellung eines Zeichens meist nie ausreicht, müssen alle Zeichen die hierbei dargestellt werden sollen in einen Zusammenhang gebracht werden. Die Überführung in diesen Zusammenhang wird hier als Synthese bezeichnet.

## 5. Integration des Coping Algorithmus

Diese Arbeit soll die Grundlagen kognitiver Fähigkeiten aufzeigen und diese dann an einer praktischen Umsetzung verdeutlichen. Die Grundlagen, die in dieser Arbeit erarbeitet wurden, bieten Ansätze um eine praktische Realisierung eines intelligenten Systems zu vollziehen. Dazu ist es notwendig, ein System zu entwickeln, das gewonnene Signale bzw. Informationen deutet und ihre Bedeutung nutzt um daraus Verhalten zu erlernen.

Hierbei werden die psychologischen Ansätze aufgezeigt, und die Fähigkeit zum Lernen mathematisch formuliert und dann in MATLAB praktisch realisiert.

Diese mathematischen Formulierungen sollen die Grundcharakteristik eines intelligenten Systems darstellen.

Dies wurde bereits von M. Sc. Werner Meyer in seiner Masterthesis realisiert und in dieser Arbeit erneut aufgegriffen.

Das Coping wurde auf eine neue Art realisiert und in der vorgegebenen Simulationssoftware implementiert. Um die Effektivität des neuen Algorithmus zu bewerten wurden Messungen durchgeführt, die die Anzahl der Lernschritte messen.

Um die Aktionen darzustellen, über die der Agent verfügt nach dem Anwenden des Coping Verfahren, wurde die Visualisierung in Form einer Merkmal-Werte-Relation implementiert. Diese Relation stellt, wie in Abbildung 7, die semantische Datenstruktur für die Aktionen sowie für die Zustände dar.

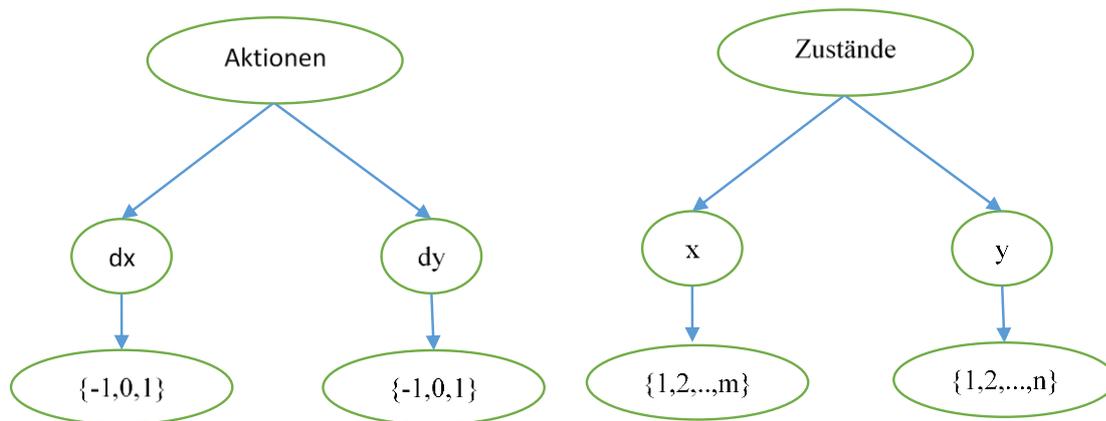


Abbildung 7: Darstellung semantischer Datenstrukturen

Für die Implementierung ist es nun wichtig, dass aus dieser semantischen Datenstruktur neue Aktionen generiert werden. Standardmäßig kennt der Agent, der im nächsten Kapitel genauer beschrieben wird, nur zwei Bewegungsrichtungen. In x und in y Richtung. Er würde also im kartesischen Koordinatensystem die Bewegungsrichtungen  $[1,0]$  und  $[-1,0]$  für das Laufen nach links und rechts von Anfang an kennen, sowie die Richtungen  $[0,1]$  und  $[0,-1]$  für die Bewegung nach oben und unten.

Der Agent kann nun diese Bewegungen neu kombinieren beispielsweise zu  $[1,1]$  für eine diagonale Bewegung. Wie bereits beschrieben verfügen finale Systeme über ein inneres Modell ihrer Umgebung. Diese ist in der Software als eine Matrix gespeichert. Um zu prüfen ob diese neu gewonnene Aktion hilfreich ist, führt er eine Probehandlung innerhalb dieser Matrix durch und bewertet das Ergebnis danach. Ist das Ergebnis zielführend, wird die Aktionsmatrix um diese Aktion erweitert. Wenn nicht, wird sie gelöscht und eine neue Kombination durchgeführt. Diese Probehandlung hat den evolutionären Vorteil, dass neue Aktionen erst geprüft werden können, bevor sie tatsächlich ausgeführt werden. Das erhöht die Sicherheit für den Agenten enorm, da gefährliche Zustände schon im vornherein vermieden werden können.

Die semantische Datenstruktur kann beliebig erweitert werden um eine neue Raumrichtung oder um einen höheren Wert des Zustandswechsels, wie z.B. in den übernächsten Zustand sich begeben anstatt nur in den nächsten.

## 6. Experimentiersystem: kognitive Maus „Theseus“

Das Experimentiersystem kognitive Maus „Theseus“ wurde von Claude Shannon erfunden und bestand ursprünglich aus einfachen Schaltkreisen und mechanischen Komponenten.

Sie hatte die Aufgabe, dass sie in ein beliebiges Labyrinth gesetzt werden konnte und sie das Labyrinth erlernen konnte. Egal wo die Maus hingestellt wurde, sie hat immer zu ihrem Ziel gefunden. Diese Lernfähigkeit ist ein wichtiger Teil des kognitiven Apparates.

In dem vorgebendem Experimentiersystem ist die Maus nicht mechanisch umgesetzt, sondern als Simulation im Programm MATLAB. Hierbei wird ein beliebiges Labyrinth vorgegeben, das nach dem Belohnungsprinzip erlernt wird. Der Agent orientiert sich im Labyrinth nun nach den maximalen Belohnungen.

### 6.1 Der kognitive Kreis an der Theseuschen Maus

Der Agent erlernt wie in Kapitel drei beschrieben das Labyrinth in dem er Belohnungswerte für jeden Zustand verteilt. Das Lernen findet praktisch gesehen so statt, dass der Agent in einen zufälligen Zustand gesetzt wird. In diesem Zustand führt er eine zufällige Aktion aus und nimmt den maximalen Wert an, der sich in diesem Zustand befindet. Der Agent besucht auf diese Art jeden Zustand und bewertet für seine möglichen Aktionen den Erfolg, dass er näher an sein Ziel kommt.

In Abbildung 8 und 9 wird ein Labyrinth aufgezeigt, an dem die Simulationen durchgeführt worden. Die dunkelblauen Felder sind die Zustände, die der Agent betreten kann. Die braunen Felder sind Hindernisse und demnach Zustände die der Agent nicht betreten kann.

Innerhalb seiner Umgebung kann sich der Agent nach der Verteilung der maximalen Belohnung zum Ziel hin orientieren.

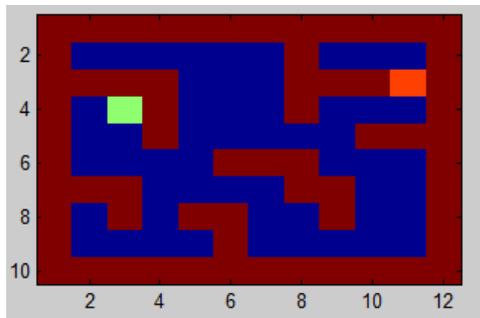


Abbildung 8: Ausgangslage

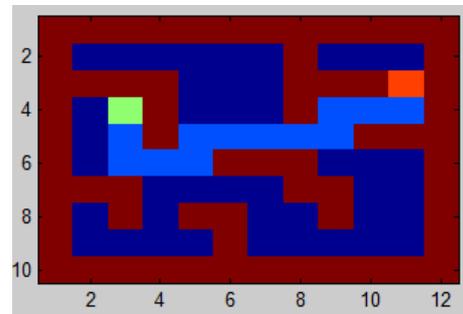


Abbildung 9: effektiv gefundene Route

In Abbildung 9 wird die gefundene Route durch die hellblau markierten Zustände markiert.

Nach dem Lernen der Umgebung kennt von jeder Stelle aus der Agent die effektive Route zum Ziel. Egal wohin der Agent platziert wird, er findet von jeder Stelle aus sein Ziel, da für jeden Zustand Belohnungswerte verteilt sind.

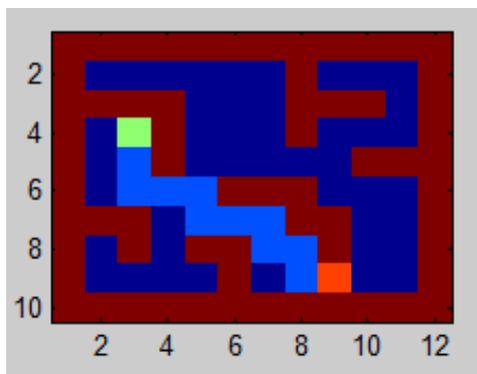


Abbildung 10: neuer Startpunkt 1

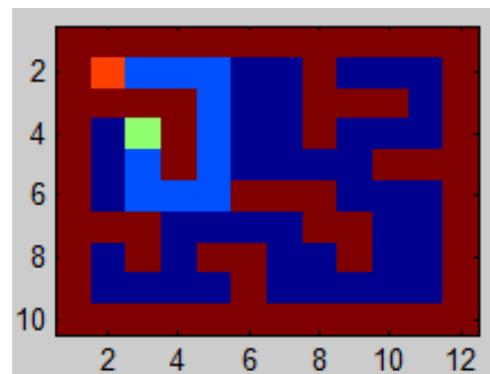


Abbildung 11: neuer Startpunkt 2

## 6.2 Messung des Lernaufwands

Um eine Vorstellung zu bekommen wie viele Schritte notwendig sind damit das Labyrinth eindeutig gelernt wurde, wird zum Lernen eine Messung durchgeführt, die im Mittel die Abweichung der alten Belohnung der Aktion im Zustand zu der neuen ermittelten Belohnung erfasst, wie in Kapitel drei beschrieben. Da am

Anfang alle Belohnungen bis auf das Ziel auf null gesetzt sind, ist auch die mittlere Abweichung anfangs null. Sobald der Agent aber im Lernen auf einen Zustand nahe des Ziels kommt, bekommt die Aktion, die zum Ziel führt, eine maximale Belohnung von 1000. So verteilen sich die Belohnungen immer weiter um den Zustandsraum herum und die Belohnungsdifferenzen minimieren sich immer weiter, bis sie gegen null konvergieren. Mit dieser Messung kann betrachtet werden wie viele Schritte notwendig sind, damit das Labyrinth sauber erlernt wurde.

Die Lernkurve wurde in Abbildung 12 wurde für über 50 Belohnungsdifferenzen gemittelt dargestellt.

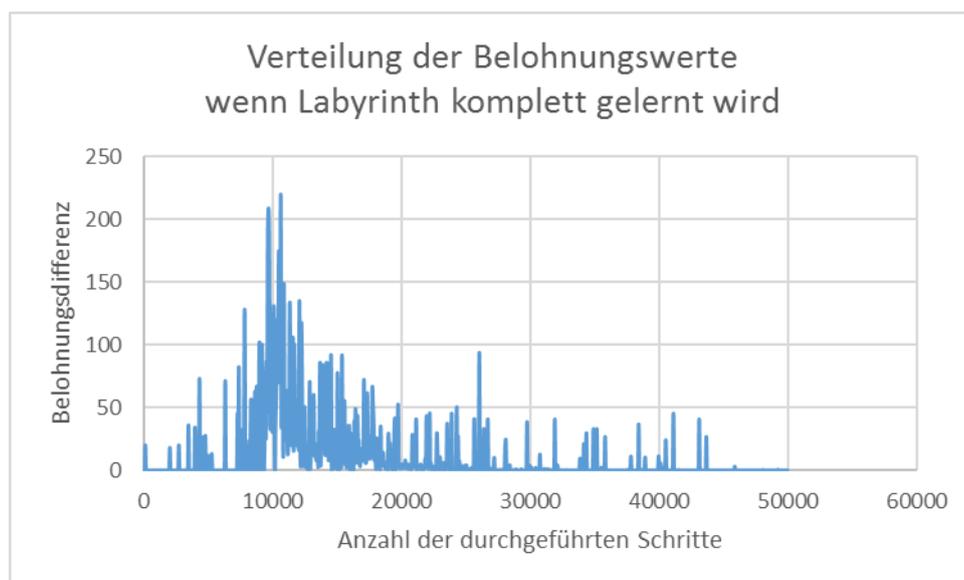


Abbildung 12: Lernverlauf bei komplett Lernen des Labyrinths

Zu erkennen ist, dass die Lernkurve anfangs steigt und ihr Maximum bei etwa 10000 Schritten erreicht. In diesem Bereich werden die Belohnungswerte am stärksten neu verteilt. Ab 30000 Schritten werden die Belohnungswerte nur noch geringfügig neu verteilt, sodass man sagen kann, dass das Lernen hier schon erfolgreich war. Komplettausgelernt hat der Agent das Labyrinth nach 45000 Schritten aber erst vollständig.

Nach dem Prinzip des Q-Lernens werden zufällig Zustände besucht und durch eine zufällige Aktion der verminderte Belohnungswert des Folgezustands angenommen. Wird nun ein Hindernis dem Agenten in den Weg gesetzt, sodass er eine Adaption durchführen muss, ist der Aufwand des Lernens logischerweise wesentlich geringer, da sich nur die Belohnungen um das Hindernis herum ändern müssen. In der Abbildung 14 ist ein Hindernis an der Stelle (8;5) gesetzt worden was der Agent nicht überwinden kann und das Labyrinth neu lernen muss. Abbildung 13 zeigt den dafür benötigten Lernaufwand.

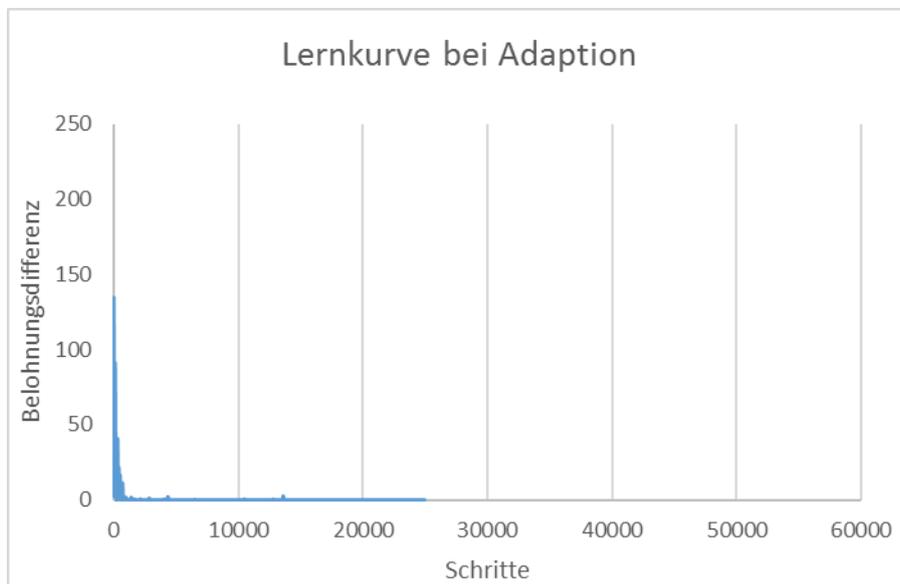


Abbildung 13: Lernkurve bei Adaption

Der Lernaufwand ist minimal, es werden etwa 200 Schritte benötigt, damit das Labyrinth mit dem Hindernis neu erlernt werden kann um wieder einen optimalen Weg zu finden.

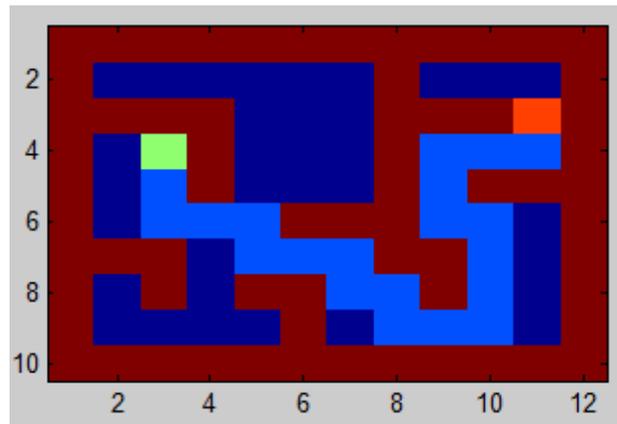


Abbildung 14: gefundene Route bei Adaption

## 6.2 Implementierung des Coping

Manchmal ist es nicht die optimale Variante das gesamte Labyrinth neu zu lernen. In Abbildung 15 wird dies verdeutlicht. Die Adaption würde dem hellblauen Pfad ermitteln, was ein viel längerer Weg ist als es nötig wäre um das Ziel zu erreichen. Der weiße Pfad verdeutlicht den Alternativweg der effektiver zum Ziel führen würde. Dazu vollzieht der Agent nun Probehandlungen, die auf den bereits bekannten Handlungen basiert.

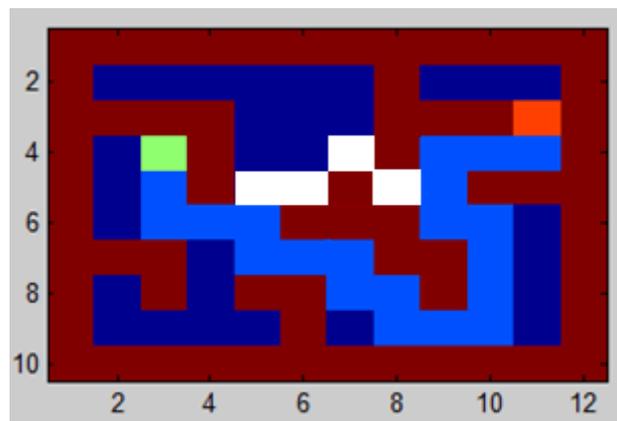


Abbildung 15: Adaption trotz alternativen Weg

Da der Agent allerdings nur seine Standardaktionen kennt, muss er um das Hindernis zu bewältigen aus diesen Aktionen eine neue Aktion ermitteln, da mit seinem Repertoire der Agent das Hindernis nicht überwinden kann.

Da die Karte zweidimensional ist gibt es also nur zwei Bewegungsrichtungen, Bewegungen in x Richtung (Ost, West) und y Richtung (Nord, Süd). Um das Hindernis wie in Abbildung 15 zu überwinden muss er eine Aktion finden in x und y Richtung mit der er das Hindernis neu überwinden könnte.

Dies sei an dem Szenario von Abbildung 15 noch einmal erklärt. Befindet sich der Agent im Zustand (8;5) stellt er ein Hindernis auf (7;5) fest. Seine Bewegung wäre vom aktuellen Zustand aus Westen also  $a = (-1,0)$ . Nun betrachtet der Agent, in welche Richtung er vom den nächsten Zustand, wo jetzt das Hindernis ist, gelaufen wäre. Diese Richtung wäre ebenso Westen, also ergäbe sich als neue Aktion:  $a_{neu} = (-2; 0)$ . Da „-2“ aber bedeuten würde, dass der Agent ein Feld überspringen würde, dies aber laut der Aktionsmatrix nicht möglich ist, ist diese neue Aktion ungültig. Nun wird die Aktion aus dem Folgezustand gewählt, die die zweithöchste Belohnung hat. Dies wäre die Aktion Norden  $a = (0,1)$  was die neue Aktion Nordwesten  $a_{neu} = (-1,1)$  bedeuten würde. Da beide Argumente der neuen Aktion auch Bestandteil der Aktionsmatrix sind, ist diese Aktion auch gültig.

Das Labyrinth muss demnach nicht neu gelernt werden, da Belohnungswerte für jeden Zustand schon existieren. Die Belohnungswerte für die neue Aktion müssen nur umverteilt werden. Dazu durchläuft der Agent in seiner inneren Welt, also in seiner Q-Matrix einmal jeden Zustand und weist der neuen Aktion in dem Zustand die maximale Belohnung aus dem Zustand zu in dem die neue Aktion führen würde. Diese Umverteilung benötigt nur so viele Schritte wie Zustände vorhanden sind. Bei dem gegebenen Labyrinth würden es  $n \cdot m$  Schritte brauchen, wobei n und m Länge und Breite des Labyrinths sind.

Nach dem Verteilen der Belohnungswerte kann der Agent das Labyrinth auf dem kürzesten Weg neu bewältigen, da er nun sein Aktionsrepertoire erweitert hat. In Abbildung 16 wird der neue effektive Weg dargestellt.

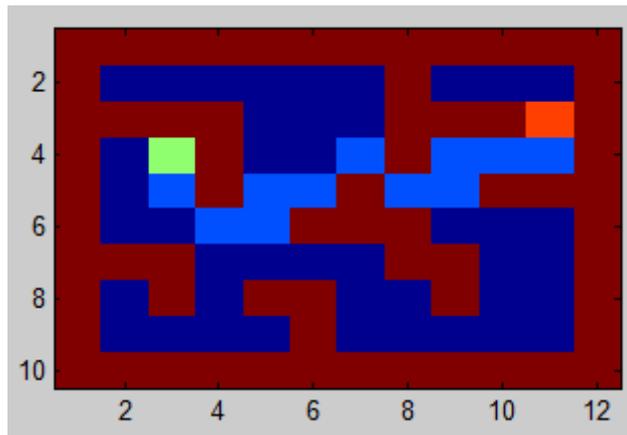


Abbildung 16: kürzester Weg mit neuen Aktionen

Um vergleichen zu können wie effektiv der neue Algorithmus ist im Vergleich zum Prinzip des Umverteilen, wurde nachdem Hinzufügen einer neuen Aktion das Labyrinth noch einmal gelernt. Das Ergebnis ist in Abbildung 17 zu sehen.

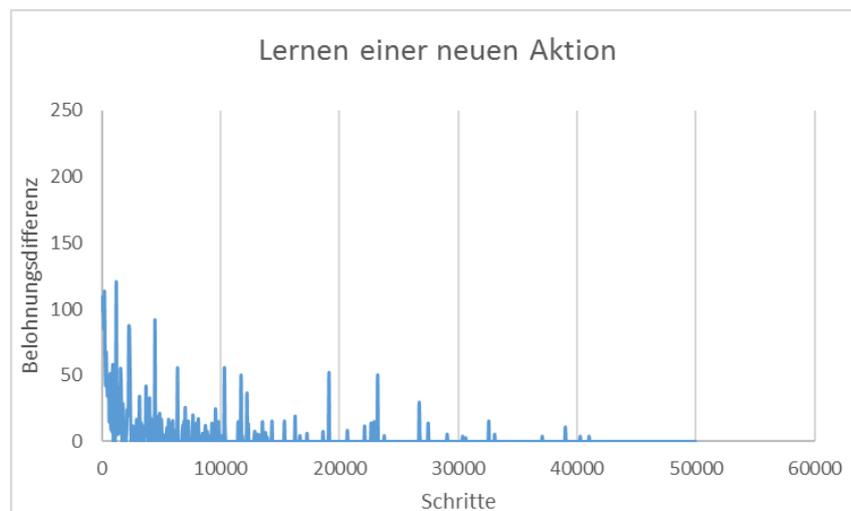


Abbildung 17: Lernaufwand für eine neue Aktion

Das Labyrinth selber weist lediglich 99 Zustände auf, die der Agent betreten kann. Anhand dieser 99 Zustände benötigt der Agent um die 10.000 Schritte um die neue Aktion neu zu erlernen. Mit dem Prinzip des Umverteilen werden lediglich 99 Schritte benötigt.

## 6.4 Visualisierung der Merkmal-Werte-Relation

Um zu erkennen über welche Aktionen der Agent verfügt, wird eine sogenannte Merkmal-Werte-Relation genutzt. In dieser wird in Form eines Baumdiagramms die Merkmale des Agenten dargestellt, und welche Werte diese Merkmale haben kann. Wie in Abbildung 18 dargestellt besitzt der Agent lediglich die Aktionen Horizontal und Vertikal und dabei die Werte von -1, 0 und 1, was bedeuten würde rückwärtslaufen, nichts tun oder vorwärts gehen. Ordnet man die Merkmale einander zu lässt sich daraus eine Richtung wie bei einem Kompass interpretieren

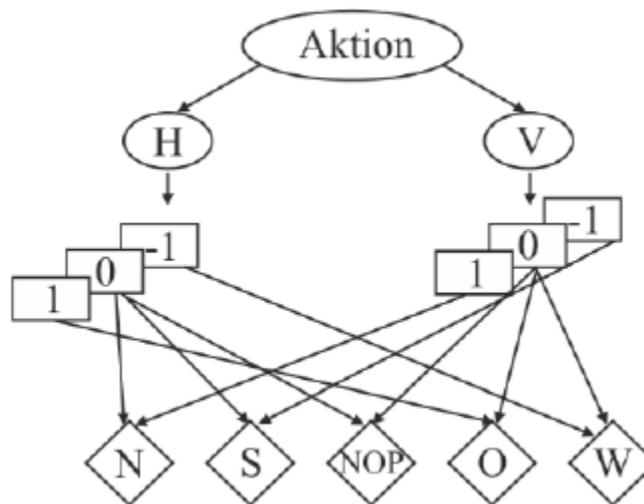


Abbildung 18: klassische Merkmal-Werte-Relation [RÖM15]

Der Agent verfügt bereits von Anfang an über seine Aktionsmatrix, in der seine bekannten Aktionen gespeichert sind. Findet der Agent durch Coping eine neue Aktion wurde diese der Aktionsmatrix hinzugefügt und dem Merkmal-Werte-Relation wurde ein weiterer Pfad hinzugefügt. In Abbildung 19 ist die ursprüngliche Merkmal-Werte Relation mit den Standardaktionen des Agenten dargestellt.

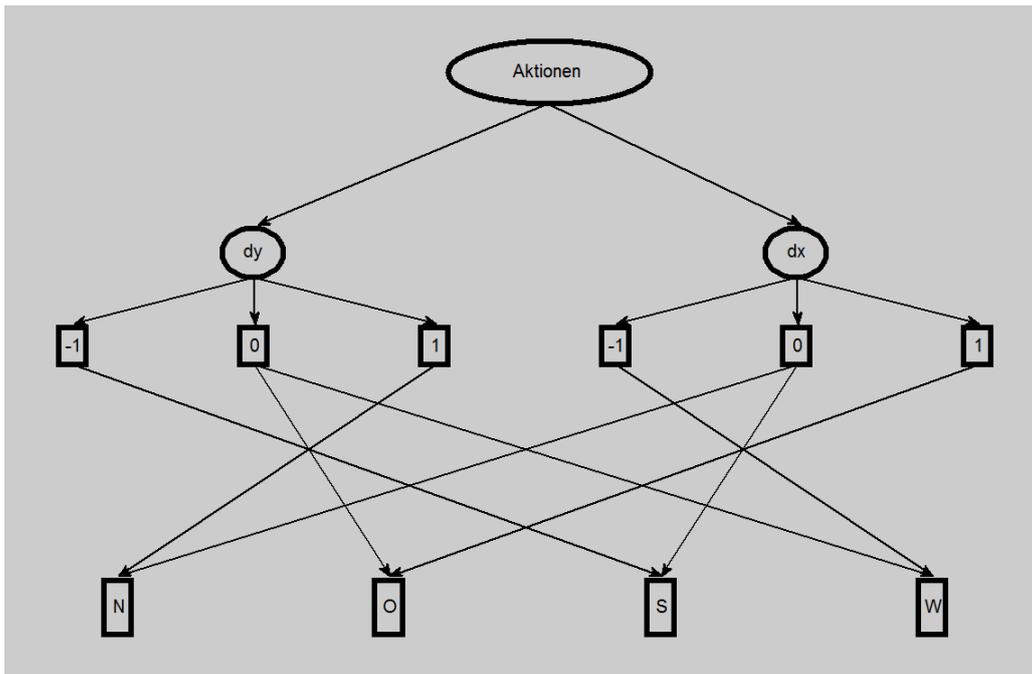


Abbildung 19: Visualisierung im Programm

Abbildung 20 zeigt die neue Merkmal-Werte Relation des Agenten nachdem er eine neue Aktion erlernt hat. Die Visualisierung fügt aus den Werten für die Merkmale automatisch die Verbindung zu der Aktion hinzu und bezeichnet sie auch von selbst durch Kombination der alten Bezeichnungen. So wird aus der Aktion Süd (-1,0) und aus der Aktion West (0,-1) die neue Aktion Nordwest(-1,-1).

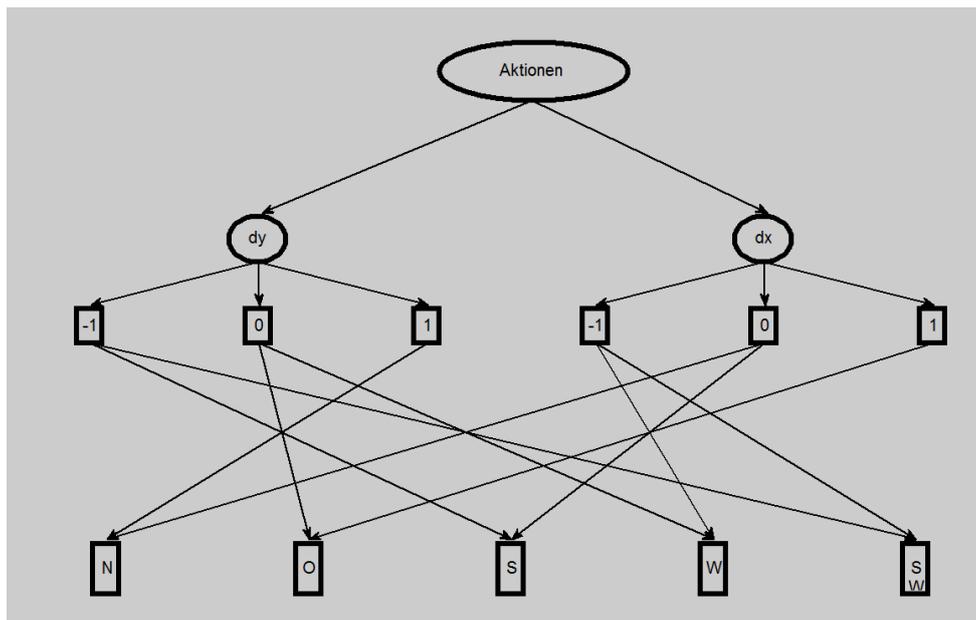


Abbildung 20: MWR nach Erlernen einer neuen Aktion

Die Visualisierung der Merkmal-Werte-Relation passt sich den Möglichkeiten des Agenten an. Würde der Agent noch eine Bewegung in z-Richtung durchführen können, wird der Visualisierung automatisch ein weiteres Merkmal hinzugefügt. Ebenso kann der Agent über mehrere Werte verfügen, wie z.B. 2 und -2. Dies würde bedeuten, dass der Agent einen Zustand „überspringen kann“. Auch hier passt sich die Visualisierung von selbst darauf an um alle Merkmale darzustellen, wie in Abbildung 21.

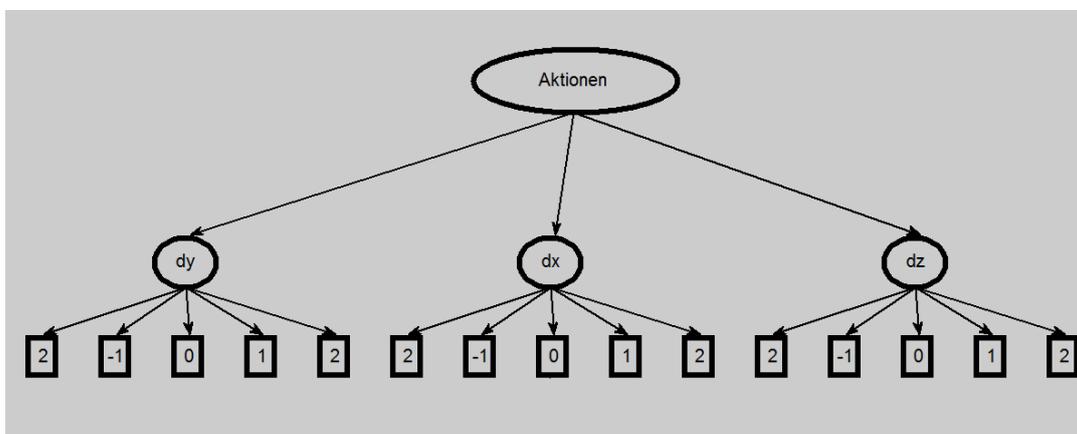


Abbildung 21: MWR mit mehreren Merkmalen und Werten

Eine weitere Anzeige ermöglicht es den aktuellen Zustand des Agenten anzeigen zu lassen, mit einer ob er das Ziel erreicht hat. Abbildung 22 zeigt wie die Darstellung dieser Anzeige implementiert wurde.

Dabei gibt POS den aktuellen Zustand an in dem sich der Agent befindet und REV die Information ob das Ziel erreicht wurde. Diese Darstellung aktualisiert sich mit jedem Zustand den der Agent betrifft.

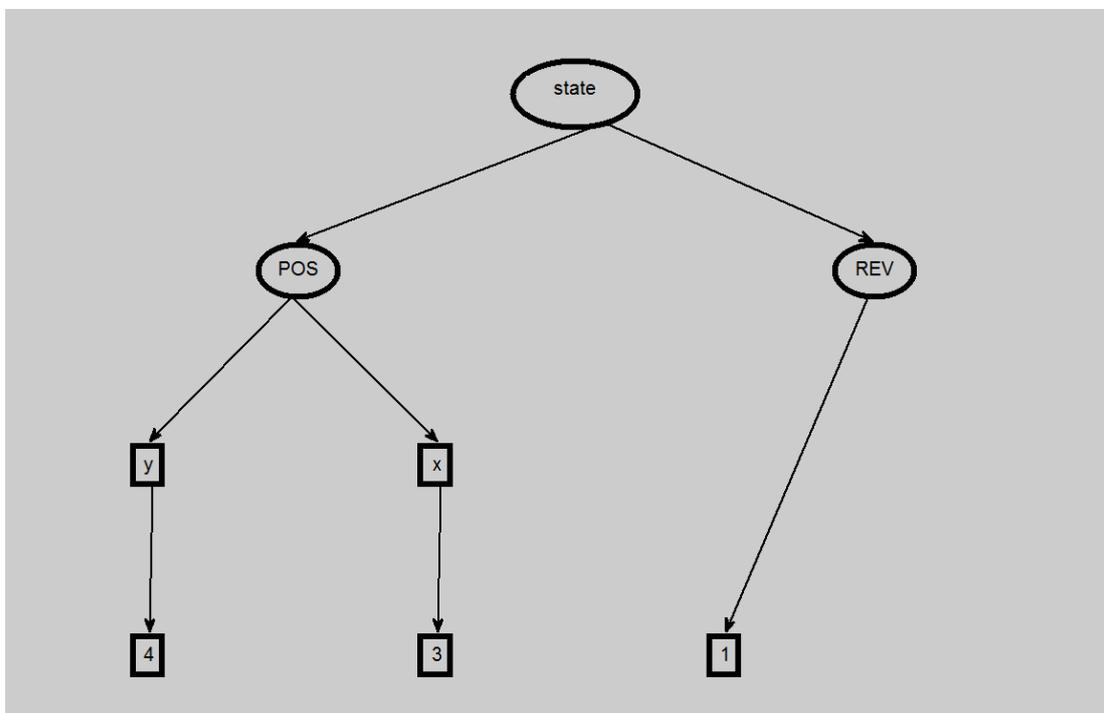


Abbildung 22: State-Anzeige des Agenten

## 7. Fazit

Diese Arbeit liefert eine Variante zur Implementierung eines lehrfähigen und anpassungsfähigen Systems. Dabei wurden auf psychologischen Grundlagen der Handlung von Organismen zurückgegriffen um diese in mathematischen Prinzipien zu formulieren.

Als effektivster Ansatz für ein intelligentes System hat sich gezeigt, dass der Konstruktivismus eine gute Basis ist, auf deren Aspekte man zurückgreifen kann in der praktischen Umsetzung. Das Lernen eines kognitiven Organismus wird von den Erfahrungen geprägt die es macht. Positive Erfahrungen werden als Belohnung angesehen, negative als Bestrafung. Somit richtet sich der Organismus anhand seiner Erfahrungen immer zu den Aktionen, für die er belohnt wird.

In der praktischen Umsetzung wird auf dieses Prinzip ebenso zurückgegriffen, in der Form, dass Zustände, die zum Ziel hinführen mit hohen Werten belegt werden, Zustände, die vom Ziel wegführen mit niedrigen.

Die erarbeiteten Prinzipien wurden in dem Programm MATLAB implementiert und an dem Beispiel der Theseuschen Maus. Der Agent hat über das Q-Lernen seine Umgebung erforscht und so eine Strategie entwickeln können um sein Ziel zu erreichen.

Der neu entwickelte Coping Algorithmus hat sich als effektiver erwiesen, als ein komplettes neu lernen des Labyrinthes, da für das Q- Lernen der neuen Aktion des Labyrinthes auf mehrere tausend Schritte zurückgegriffen werden müsste, jetzt jedoch nur noch so viele Schritte benötigt werden wie es Zustände gibt.

Die verfügbaren Aktionen wurden als grafische Ausgabe in Form einer Merkmal-Werte-Relation implementiert, sodass man sehen kann ob und welche Aktionen der Agent gelernt hat und aus welchen Werten sie sich zusammensetzt. Ebenso wurde die Möglichkeit einer Zustandsanzeige implementiert, sodass man während des gesamten Durchlaufen des Labyrinthes sehen kann, in welchem Zustand sich der Agent befindet und ob er das Ziel schon erreicht hat.

Da die Visualisierung sich selber anpasst an neue Gegebenheiten, könnte nachfolgend das Experimentiersystem um eine z-Richtung erweitert werden und so ein dreidimensionales Labyrinth ermöglichen.

Ebenso könnten die Möglichkeiten des Agenten noch um Aktionen wie „Felder überspringen“ erweitert werden, was ein noch höheres Aktionsrepertoire des Agenten ermöglichen würde.

## Literaturverzeichnis

- [BIS09] Bischof, N. (2009). *Psychologie - Ein Grundkurs für Anspruchsvolle*. Kohlhammer.
- [ERT09] Ertel, W. (2009). *Grundkurs Künstliche Intelligenz*. Vieweg + Teubner.
- [GES12] *Gesundheit.de*. (24. April 2012). Von <http://www.gesundheit.de/wissen/haetten-sie-es-gewusst/persoenlichkeiten-der-medizin/pawlow-der-entdecker-des-bedingten-reflexes> abgerufen
- [MEY15] Meyer, W. (2015). *Modellierung und Implementierung einer kognitiven Verhaltenssteuerung*. Cottbus.
- [ROB12] Robak, S. (2012). *Was ist Konstruktivismus*. München: GRIN Verlag.
- [RÖM13] Ronald Römer, G. W. (2013). *Ein Beitrag zu den Natur – und Geisteswissenschaftlichen Grundlagen kognitiver Systeme*. BTU Cottbus.
- [RÖM15] Ronald Römer, M. W. (2015). *Konzeption eines kognitiven Systems für den experimentellen Einsatz in Forschung und Lehre*. BTU Cottbus.
- [RUS04] Russel, S. (2004). *Künstliche Intelligenz ein moderner Ansatz*. Pearson Studium.
- [THO16] Thompson, B. Abgerufen am 18. Juni 2016 von <http://www.wissen-hund.de/desensibilisierung/>
- [WAL04] Walter, D. O. (2004). Abgerufen am 12. Juli 2016 von [http://www.verhaltenswissenschaft.de/Psychologie/Behaviorismus/Methodologischer\\_Behaviorismus/Klassischer\\_Behaviorismus/Watson/Behavioristisches\\_Manifest/behavioristisches\\_manifest.htm](http://www.verhaltenswissenschaft.de/Psychologie/Behaviorismus/Methodologischer_Behaviorismus/Klassischer_Behaviorismus/Watson/Behavioristisches_Manifest/behavioristisches_manifest.htm)

## Abbildungsverzeichnis

Abbildung 1: Darstellung innere gegen äußere Sicht [BIS09] .....	10
Abbildung 2: Entwicklungspyramide [BIS09].....	11
Abbildung 3: Relation zwischen Umgebung und Agenten .....	14
Abbildung 4: Kognition nach Heykins .....	16
Abbildung 5: Verlauf der Lernkurve des Agenten [MEY15] .....	26
Abbildung 6: Feinstruktur eines kognitiven Agenten [RÖM13].....	28
Abbildung 7: Darstellung semantischer Datenstrukturen .....	31
Abbildung 8:Ausgangslage .....	33
Abbildung 9:effektiv gefundene Route .....	33
Abbildung 10: neuer Startpunkt 1.....	33
Abbildung 11: neuer Startpunkt 2.....	33
Abbildung 12: Lernverlauf bei komplett Lernen des Labyrinths .....	34
Abbildung 13: Lernkurve bei Adaption .....	35
Abbildung 14: gefundene Route bei Adaption.....	36
Abbildung 15: Adaption trotz alternativen Weg.....	36
Abbildung 16: kürzester Weg mit neuen Aktionen .....	38
Abbildung 17: Lernaufwand für eine neue Aktion.....	38
Abbildung 18:klassische Merkmal-Werte-Relation [RÖM15].....	39
Abbildung 19: Visualisierung im Programm.....	40
Abbildung 20: MWR nach erlernen einer neuen Aktion.....	41
Abbildung 21: MWR mit mehreren Merkmalen und Werten .....	41
Abbildung 22: State-Anzeige des Agenten.....	42

## Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass ich die vorliegende BA/MA-Arbeit selbstständig und ohne fremde Hilfe verfasst, andere als die angegebenen Quellen und Hilfsmittel nicht benutzt bzw. die wörtlich oder sinngemäß entnommenen Stellen als solche kenntlich gemacht habe. Mit meiner Unterschrift stimme ich einer Veröffentlichung der Arbeit oder Teilen aus der Arbeit zu.

---

Daniel Götze