

**Beiträge des Instituts für Umweltsystemforschung  
der Universität Osnabrück**

**Herausgeber: Prof. Dr. Michael Matthies**

**Beitrag Nr. 36**

**Agentenbasierte Modellierung von  
Persönlichkeitsunterschieden auf der Grundlage  
der PSI-Theorie**

**Ivo Siekmann**

**Juli 2006**



Institut für  
Umweltsystemforschung  
Universität Osnabrück

**ISSN-Nr. 1433-3805**

**Beiträge des Instituts für Umweltsystemforschung  
der Universität Osnabrück**

**ISSN 1433-3805**

Herausgeber

Prof. Dr. Michael Matthies  
Universität Osnabrück  
Institut für Umweltsystemforschung  
Artilleriestr. 34

D-49069 Osnabrück

Tel. 0541/969-2575  
Fax. 0541/969-2599

E-Mail: [matthies@uos.de](mailto:matthies@uos.de)  
<http://www.usf.uni-osnabrueck.de>

© USF – Institut für Umweltsystemforschung, Universität Osnabrück

---

## WHITE RABBIT

ONE PILL MAKES YOU LARGER, AND ONE PILL MAKES YOU SMALL  
AND THE ONES THAT MOTHER GIVES YOU, DON'T DO ANYTHING AT ALL;  
GO ASK ALICE, WHEN SHE'S TEN FEET TALL.

AND IF YOU GO CHASING RABBITS, AND YOU KNOW YOU'RE GOING TO FALL;  
TELL 'EM A HOOKAH-SMOKING CATERPILLAR HAS GIVEN YOU THE CALL;  
TO CALL ALICE, WHEN SHE WAS JUST SMALL.

WHEN THE MEN ON THE CHESSBOARD GET UP AND TELL YOU WHERE TO GO;  
AND YOU'VE JUST HAD SOME KIND OF MUSHROOM, AND YOUR MIND IS MOVING LOW;  
GO ASK ALICE, I THINK SHE'LL KNOW.

WHEN LOGIC AND PROPORTION HAVE FALLEN SLOPPY DEAD;  
AND THE WHITE KNIGHT IS TALKING BACKWARDS;  
AND THE RED QUEEN'S OFF WITH HER HEAD;  
REMEMBER WHAT THE DORMOUSE SAID,  
FEED YOUR HEAD, FEED YOUR HEAD

Jefferson Airplane: "White Rabbit", 1967, [7]

VINCENT

Yes, but do you consider a dog to  
be a filthy animal?

JULES

I wouldn't go so far as to call a  
dog filthy, but they're definitely  
dirty. But a dog's got  
personality. And personality goes  
a long way.

VINCENT

So by that rationale, if a pig had  
a better personality, he's cease to  
be a filthy animal?

JULES

We'd have to be talkin' 'bout one  
motherfuckin' charmin' pig. It'd  
have to be the Cary Grant of pigs.

Quentin Tarantino: "Pulp Fiction", 1994, [16]

# Danksagung

Diese Arbeit war auf Unterstützung durch viele Menschen angewiesen. Ihnen allen bin ich zu großem Dank verpflichtet.

Zuallererst gilt mein Dank meinen beiden Betreuern Claudia Pahl-Wostl und Julius Kuhl: Zunächst dafür, dass sie es mir ermöglicht haben, über ein wenig alltägliches Thema zu schreiben; vor allem jedoch für die freundliche und ermutigende Betreuung, durch die ich auch in teilweise entmutigenden Phasen die nötige Motivation bekam.

Ohne die fachliche Unterstützung von Marc Strickert und Stephan Jätzold wäre ich über die ersten paar Seiten nicht hinausgekommen: Marc Strickert gab mir die nötige Orientierungshilfe, um sicher durch den beinahe undurchdringlichen Dschungel der neuronalen Netzwerkmodelle zu gelangen. Von Stephan Jätzold habe ich die entscheidende Starthilfe für den Umgang mit der neuen Quicksilver-Version erhalten.

Nicht im Zusammenhang mit dieser Arbeit, sondern für mein ganzes Studium möchte ich mich besonders bedanken bei Eberhard Umbach: Er nahm sich — wie für viele andere Studierende — auch für mich sehr viel Zeit. Ich werde mich immer gerne erinnern an lange, hochinteressante Gespräche über beinahe jedes vorstellbare Thema.

Anja Kipp danke ich für die — während meines gesamten Studiums andauernde — Abschirmung von universitärer Bürokratie. Weil ich zumindest davon gehört habe, dass es an manchen Universitäten schwieriger ist, sich für eine Prüfung anzumelden als sie zu bestehen, gehe ich davon aus, dass Anja Kipp diejenige ist, die die bürokratischen Hürden auf wundersame Weise zum Verschwinden bringt.

Stellvertretend für viele Freunde, Kommilitonen, Verwandte und Bekannte, denen ich mal mehr, mal weniger erfolgreich versuchte, zu erklären, mit was ich mich in der letzten Zeit genau beschäftigt habe, möchte ich Angelika und Andreas Siekmann, meine Eltern, nennen, die einen der erfolgloseren Erklärungsversuche zunächst geduldig ertrugen — um mir dann zu sagen, wie ich es besser machen kann.

Meiner Freundin Tamara Rarău danke ich schließlich dafür, dass sie mich immer wieder aus dem “Wunderland” in die wirkliche Welt zurückholte.

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>8</b>
<b>2</b>	<b>Ein Modell für Objekterkennung und Verhaltenssteuerung</b>	<b>9</b>
1.	Das Objekterkennungssystem . . . . .	9
1.1	Objekterkennung als Kategorisierung . . . . .	10
1.2	Darstellung von Kategorien als Bitstrings . . . . .	10
1.3	Hervorhebung von Unterschieden . . . . .	10
2.	Bedürfnisse . . . . .	11
2.1	Bedürfnisbefriedigung durch bestimmte Kategorien . . . . .	11
2.2	Ist-Sollwert-Modell der Bedürfnisbefriedigung . . . . .	12
2.3	Situationsabhängiges Verhalten . . . . .	13
2.4	Wahrgenommene und tatsächliche Befriedigung . . . . .	13
3.	Die intuitive Verhaltenssteuerung (IVS) . . . . .	13
3.1	Basisverhaltensweisen . . . . .	14
3.2	Beispiel für “unscharfe” WENN-DANN-Regeln . . . . .	14
3.3	Verhaltensprogramme . . . . .	15
3.4	Verhaltensprogramme und Planen . . . . .	16
4.	Affektdynamik . . . . .	17
4.1	Differenzgleichungsmodell der Affektdynamik . . . . .	19
4.2	Korrekturen des Differenzgleichungsmodells . . . . .	19
4.3	Bedürfnisgenerierter und antizipierter Affekt . . . . .	20
4.4	Parametrisierung . . . . .	20
<b>3</b>	<b>Ein Modell für Extensionsgedächtnis, Fühlen und Selbst</b>	<b>22</b>
1.	Fühlen, Selbst und Extensionsgedächtnis in der PSI-Theorie . . . . .	22
1.1	Das Fühlen . . . . .	22
1.2	Das Selbst . . . . .	23
1.3	Kongruenzorientierte Aufmerksamkeit (Vigilanz) . . . . .	25
1.4	Das Extensionsgedächtnis (EG) . . . . .	25
1.5	Fühlen als Zusammenspiel von OES und EG . . . . .	26
2.	Neuronale Netzwerkmodelle . . . . .	31
2.1	Allgemeine Eigenschaften . . . . .	31
2.2	Feedforward-Netze als Funktionsapproximatoren . . . . .	32
2.3	Hopfield-Netze als Assoziativspeicher . . . . .	33
2.4	Ähnlichkeit von Assoziationen . . . . .	38
3.	Modellbeschreibung . . . . .	41
3.1	Das OES als Lieferant von Eindrücken . . . . .	41

---

3.2	Der Hippocampus . . . . .	41
3.3	Das EG als Hopfield-Netz . . . . .	46
3.4	Affekteinflüsse . . . . .	48
3.5	Verhaltenssteuerung durch das EG . . . . .	48
4.	Experiment zu den Inhalten des EG . . . . .	49
4.1	Versuchsbeschreibung . . . . .	50
4.2	Beobachtungen . . . . .	51
4.3	Auswertung . . . . .	53
4.4	Theoretische Überlegungen zu den Ergebnissen des Simulationsexperiments . . . . .	59
<b>4</b>	<b>Fazit</b>	<b>63</b>
<b>A</b>	<b>Resultate des Simulationsexperiments</b>	<b>64</b>
1.	Einzelne im EG gespeicherte Assoziationen . . . . .	64
2.	Kombinationen der im EG gespeicherten Assoziationen . . . . .	64

---

---

# Kapitel 1

## Einleitung

Die Frage, was Persönlichkeit ist und wer sie hat, ist nicht ganz einfach — das zeigt schon der oben zitierte Dialog aus Quentin Tarantinos Film “Pulp Fiction”. Das Ansinnen, ausgerechnet ein so schwieriges, formal kaum fassbares Phänomen wie Persönlichkeitsunterschiede durch eine abstrakte Methode wie Computermodelle erforschen zu wollen, erscheint deshalb — zumindest auf den ersten Blick — vermessen. Doch die durch sogenannte *Agentenmodelle* gegebene Möglichkeit, die Bestandteile eines sozialen oder psychologischen Systems sehr detailliert abbilden zu können, eröffnet auch für die Psychologie einen neuen Zugang für den Einsatz von Computermodellen. Um allerdings mithilfe der enormen Möglichkeiten der agentenbasierten Modellierung nicht Beliebigkeit zu produzieren, bedarf es Orientierung. Diese erhält das Modell durch seine Basis, die PSI-Theorie von Julius Kuhl. Dank dieser Grundlage müssen Begriffe wie Bedürfnisse oder Affekt bei der Modellierung nicht willkürlich nach Bedarf mehr schlecht als recht entwickelt werden; stattdessen kann auf das schon bestehende Theoriegebäude zurückgegriffen werden.

Die Computersimulation “Alice im Wunderland” entstand unter Verwendung der JAVA-Bibliothek *Quicksilver* [6]. Mehrere vergleichsweise einfache Modelle, die jeweils Teilsysteme einer Person beschreiben, wurden zu einem Agenten zusammengesetzt. Ihren Namen erhielt die “Agentin” aufgrund der unverkennbaren Charaktermerkmale von Lewis Carrolls Romanfigur: Die simulierte Person erkundet wie Alice, von unstillbarer Neugier getrieben, das Wunderland, in dem sie ausgesetzt worden ist. Grundlage hierfür sind Modelle der Objekterkennung, automatischer Verhaltensroutinen und der Fähigkeit, anhand des Status von Bedürfnissen das Verhalten auszurichten (Kapitel 2).

In Kapitel 3 erhält Alice — dem Rat der Haselmaus aus “White Rabbit” folgend — die Fähigkeit, Erfahrungen zu sammeln. Im wahrsten Sinne des Wortes “füttert sie ihren Kopf” mit Informationen der Umgebung, die durch einen raffinierten Prozess ins sogenannte Extensionsgedächtnis übertragen werden. Ziel dieser Arbeit ist es, diese Art des Lernens von Erfahrungen durch die Übertragung in ein Computermodell auf einer zusätzlichen Ebene erforschbar zu machen: Das Modell erlaubt es, in den “Kopf” der simulierten Person hineinzuschauen und die dort ablaufenden Prozesse direkt zu beobachten. Es kann deshalb zur anschaulicheren Darstellung der Aussagen einer Theorie dienen, sie dadurch besser verstehen helfen und Anstöße für die Weiterentwicklung geben. Dies wird am Schluss der Arbeit anhand der Auswertung eines Simulationsexperiments demonstriert.

# Kapitel 2

## Ein Modell für Objekterkennung und Verhaltenssteuerung

Den Kern der PSI-Theorie bilden die *kognitiven Makrosysteme*; das *Objekterkennungssystem* und die *Intuitive Verhaltenssteuerung* sowie das *Planen* und das *Fühlen*. Jedes dieser vier Systeme bringt bestimmte Charakteristika der Wahrnehmung und des Verhaltens mit. Persönlichkeitsunterschiede können durch die Dominanz eines oder mehrerer Systeme erklärt werden; Wahrnehmung und Verhalten sind dann in erster Linie durch die Merkmale der dominanten Systeme charakterisiert.

Grundlegendes Konzept bei der Entwicklung dieses Modells war es, zunächst nur von der Aktivierung der vier kognitiven Systeme durch positiven und negativen Affekt auszugehen, das Modell jedoch durch Erweiterung der Systeme über die bloße Aktivierung hinaus, zu verfeinern. Das bisherige Modell konzentrierte sich in erster Linie auf das Objekterkennungssystem (OES) und die Intuitive Verhaltenssteuerung (IVS). Um elementare Wahrnehmung darstellen zu können, wurde ein einfaches Modell der durch die Objekterkennung wahrgenommenen Objekten entwickelt (Abschnitt 1.). Objekte beeinflussen die Befriedigung der *Bedürfnisse* (Abschnitt 2.): Situationsabhängig können Objekte bedürfnisbefriedigend oder schädlich wirken bzw. keinen Einfluss ausüben. Die Befriedigung der Gesamtheit aller Bedürfnisse bestimmt die Affektlage der Person; eine überwiegende Befriedigung der Bedürfnisse erzeugt positiven Affekt ( $A+$ ), sind die Bedürfnisse nicht befriedigt, entsteht negativer Affekt ( $A-$ ). Die Affektlage beeinflusst die Aktivierung der vier kognitiven Makrosysteme. Stark vereinfacht kann man diese sehr subtile *Affektdynamik* (Abschnitt 4.) so beschreiben: Bei positivem Affekt wird die Intuitive Verhaltenssteuerung aktiviert, sinkt der positive Affekt ab, wird das Planen aktiv. Negativer Affekt führt zur Aktivierung der Objekterkennung, absinkender negativer Affekt aktiviert das Fühlen. Grundlage der Affektdynamik sind die in [9, Kuhl, Kapitel 5] postulierten Modulationsannahmen.

### 1. Das Objekterkennungssystem

Elementares Wahrnehmen, d.h. die Strukturierung der Reizflut, die auf die Person aus ihrer Umgebung einströmt, ist die Aufgabe der Objekterkennung. Struktur wird durch die Auflösung der Reize in strikt voneinander abgetrennte Einzelobjekte erzeugt. Wenn nur das Verhalten einer Person unter dem Einfluss verschiedener Affekte untersucht werden soll, ist eine nähere Bestimmung dieser Einzelobjekte nicht notwendig. In diesem Fall können Objekte direkt mit

---

Affekt besetzt werden, d.h. sie lösen bei der Person positiven bzw. negativen Affekt aus. Dies ist vom psychologischen Standpunkt jedoch insofern ungenau, da Objekte allein keinen Affekt auslösen können, sondern erst dadurch, dass sie bedürfnisbefriedigend wirken oder für die Bedürfnisbefriedigung schädlich sind, von der Person mit positivem bzw. negativem Affekt bewertet werden.

## 1.1 Objekterkennung als Kategorisierung

Die Frage ist nun: Wie muss ein Objektmodell aussehen, das die Bewertung von Objekten anhand von Bedürfnissen erlaubt? Die Lösung besteht darin, dass das Modell des Objekterkennungssystems nicht einzelne Objekte erkennt, sondern Objekte in *Kategorien* einordnet. Dies entspricht noch gerade dem Verarbeitungsniveau, das Kuhl für die Objekterkennung angibt.<sup>1</sup> Durch die Einordnung in Kategorien wird die entscheidene Leistung des Objekterkennungssystems, das *Wiedererkennen* von Objekten, gut abgebildet. Denn selbst ein ganz bestimmtes, schon bekanntes Objekt wird nicht in erster Linie als einzelnes Objekt wahrgenommen, sondern in eine bestimmte, enger oder weiter gefasste Kategorie wie Person, Nachbar oder bester Freund eingeordnet (das Beispiel "bester Freund" zeigt, dass eine Kategorie durchaus nur ein einziges Objekt enthalten kann). Durch Kategorien können Verbindungen zu Bedürfnissen leicht hergestellt werden: Bedürfnisse bewerten bestimmte Kategorien situationsabhängig als bedürfnisbefriedigend oder für die Bedürfnisse schädlich.

## 1.2 Darstellung von Kategorien als Bitstrings

Das Erkennen von Objekten wird wie oben erläutert durch das Modell ausschließlich als Einordnung in Kategorien dargestellt. Dazu werden die Objekte jeweils mit einem Bitstring versehen; dieser wird verwendet, um zu überprüfen, ob das Objekt zu einer bestimmten Kategorie gehört. Kategorien werden ebenfalls durch einen Bitstring modelliert; zusätzlich wird hier durch eine positive ganze Zahl, die *Weite* der Kategorie angegeben. Durch die Weite wird festgelegt, bis zu welchem Bit von rechts gelesen, ein Objekt mit dem Bitstring der Kategorie übereinstimmen muss, um zur selben Kategorie gezählt zu werden (siehe Abb. 1.1). Es kann sehr enge Kategorien geben, zu denen nur mit einem ganz bestimmten Bitstring markierte Objekte gehören, beispielsweise nur die Objekte, die den Bitstring 11010100 tragen. Soll die Kategorie 11010100 weiter gefasst werden, werden auch diejenigen Objekte zur Kategorie 11010100 gezählt, deren Bitstrings nur bis zur fünften Stelle von rechts paarweise mit 11010100 übereinstimmen, also beispielsweise 11011101, 11010011 oder 11011111.

## 1.3 Hervorhebung von Unterschieden

Als Hauptmerkmal der Arbeitsweise des Objekterkennungssystems nennt Kuhl die Hervorhebung von Unterschieden. Diese Eigenschaft wird durch Einordnung von Objekten in Kategorien gut abgebildet: Ein Objekt gehört entweder zu einer bestimmten Kategorie oder nicht; fließende Übergänge sind nicht möglich. Ist das Objekterkennungssystem aktiviert, wird angenommen, dass es Unterschiede stärker berücksichtigt. In diesem Fall werden also sehr enge Kategorien gebildet; noch relativ ähnliche Objekte werden dann in verschiedene Kategorien eingeordnet. Ist

---

<sup>1</sup>"Persönlichkeitspsychologisch interessant sind die Endergebnisse der mit Einzelempfindungen beginnenden elementaren Wahrnehmungsprozesse. Diese Endergebnisse sind Objekte oder auf einer höheren, aber noch relativ elementaren Ebene *Kategorien*." [9, Kuhl, S.315]

---

	Übereinstimmung?	vernachlässigt
Kategorie	110101	00
getestetes Objekt	110111	11

	Übereinstimmung?	vernachlässigt
Kategorie	1101	0100
getestetes Objekt	1101	1111

Abbildung 1.1: Der unterschiedliche Umfang von Kategorien wird durch die unterschiedlich strikte Überprüfung von Bitstrings modelliert: Ein mit dem Bitstring 11011111 markiertes Objekt gehört nicht zur engen Kategorie der Objekte, die mit 11010100 von links bis zum dritten Bit von rechts übereinstimmen. Dagegen wird es zur weiter gefassten Kategorie der Objekte gezählt, die mit 11010100 bis zum fünften Bit von links übereinstimmen.

die Objekterkennung dagegen deaktiviert, werden weiter gefasste Kategorien gebildet und deshalb möglicherweise wichtige Unterschiede zwischen Objekten nicht erkannt. Im Modell wird dies durch variable Weiten der Kategorien ausgedrückt.

## 2. Bedürfnisse

Die Aufgabe der Bedürfnisse ist die *Bewertung* der aktuellen Situation aus der Sicht einer Person. Anhand der Befriedigung bzw. Nicht-Befriedigung der Bedürfnisse wird bestimmt, ob die Person positiven oder negativen Affekt empfindet. Da in vielen Fällen die Bedürfnisbefriedigung von bestimmten Objekten abhängt, werden auch die Objekte durch Bedürfnisse bewertet. Dies wird in Abschnitt 2.1 erläutert. Wie die Befriedigung der Bedürfnisse berechnet wird, ist im darauffolgenden Abschnitt 2.2 dargestellt. Abschnitt 2.3 erklärt, dass das Bedürfnismodell eingesetzt werden kann, um an die aktuelle Situation angepasstes Verhalten zu modellieren. Schließlich wird in Abschnitt 1.1 auf Ungenauigkeiten der Wahrnehmung eingegangen. Diese können dazu führen, dass die Person falsch einschätzt, ob Objekte bedürfnisbefriedigend wirken oder nicht. Die Berechnung des Affektes aus der Bedürfnislage wird später im Abschnitt über die Affektdynamik, 4.3, dargestellt.

### 2.1 Bedürfnisbefriedigung durch bestimmte Kategorien

Wie wird festgestellt, ob ein bestimmtes Objekt ein Bedürfnis befriedigt? Die Antwort lautet einfach: Wenn es zur richtigen Kategorie gehört. In diesem Modell kann das Objekt dann *antizipierten Affekt* (siehe Abschnitt 4.3) auslösen. Bedürfnisse prüfen also Objekte auf die Zugehörigkeit zu einer bestimmten Kategorie und geben dann eine Bewertung ab. Das Bedürfnis kann dabei mehr oder weniger "strikt" sein, d.h. Objekte einer engeren oder weiteren Kategorie als bedürfnisbefriedigend akzeptieren. So erscheint es beim Bedürfnis nach Nahrung in vielen Fällen angemessen — Persönlichkeitsunterschiede nicht ausgeschlossen —, anzunehmen, dass es durch eine recht weite Kategorie von Objekten befriedigt werden kann. Das Bedürfnis eines Zuckerkranken nach Insulin wird dagegen tatsächlich nur durch Insulin befriedigt — hier sollte das Bedürfnis nur eine sehr enge Kategorie von Objekten, also nur einen ganz bestimmten Bitstring akzeptieren. Die Striktheit eines Bedürfnisses, mit der nur Objekte einer eher engen oder weiteren Kategorie als bedürfnisbefriedigend akzeptiert werden, kann sich auch situationsspe-

---

zifisch ändern: Der ursprüngliche Plan einer Person, sich durch “irgendetwas im Fernsehen” unterhalten zu lassen, kann sich schon bald ändern, wenn nach einiger Zeit festgestellt wird, dass das laufende Programm leider auf allen Kanälen eher unbefriedigend ist, was bald das Bedürfnis nach einer ganz bestimmten Sendung entstehen lässt — die leider am heutigen Abend gerade nicht ausgestrahlt wird. Dies alles kann auf der Ebene der Bedürfnisse ebenso modelliert werden wie die Unterschiede in der Genauigkeit der Wahrnehmung (siehe Abschnitt 1.1) — durch unterschiedlich strikte Überprüfung der Gleichheit von Bitstrings.

## 2.2 Ist-Sollwert-Modell der Bedürfnisbefriedigung

Die Befriedigung eines Bedürfnisses ist nicht auf die beiden Extrema befriedigt bzw. unbefriedigt festgelegt: Gemäß der PSI-Theorie wird Bedürfnisbefriedigung durch Ist-Sollwert-Diskrepanzen sogenannter *Bedürfnismelder* beschrieben. Als Beispiel nennt Kuhl [9, Kuhl, S.414] als Bedürfnismelder für Hunger den Blutzuckerspiegel: Steigt der Blutzuckerspiegel über einen bestimmten Wert, fühlt man sich satt, sinkt er, entsteht ein Hungergefühl. Das Hungergefühl ist dabei um so stärker, je tiefer der Blutzuckerspiegel unter den Sollwert gesunken ist.

Die *Befriedigung* der Bedürfnisse in diesem Modell liegt zwischen den Werten 1 und -1 — positive Werte bedeuten mehr oder weniger starke Befriedigung, negative Werte die Nicht-Befriedigung des Bedürfnisses. Damit die unterschiedliche *Wichtigkeit* der Bedürfnisse berücksichtigt werden kann, ist jedes Bedürfnis mit einem Gewichtungsfaktor  $i_k$  zwischen 0 und 1 versehen. Damit wird die tatsächliche Befriedigung  $s_k$  eines Bedürfnisses durch seine Wichtigkeit  $i_k$  korrigiert:

$$r_k = s_k i_k. \quad (2.1)$$

Für die Gewichtungen  $i_k$  der Bedürfnisse muss dabei gelten

$$\sum_{k=1}^n i_k = 1. \quad (2.2)$$

Die Gesamtbefriedigung aller Bedürfnisse, die den *bedürfnisgenerierten Affekt* (siehe Abschnitt 4.3) bestimmt, wird nach folgender Formel berechnet:

$$T = \sum_{j=1}^n s_j i_j \quad (2.3)$$

Allerdings kann es nach der PSI-Theorie durchaus gleichzeitig positiven und negativen Affekt geben, so dass alternativ zwischen befriedigten und unbefriedigten Bedürfnissen differenziert werden kann:

$$S = \sum_{j=1}^n s_j^+ i_j \quad (2.4)$$

$$U = \sum_{k=1}^n s_k^- i_k \quad (2.5)$$

Hier wird getrennt betrachtet, welche Bedürfnisse befriedigt und welche unbefriedigt sind. Der positive Affekt kann dann aus dem Befriedigungsgrad  $S$ , der negative Affekt aus dem Nicht-Befriedigungsgrad  $U$  bestimmt werden.

---

## 2.3 Situationsabhängiges Verhalten

Durch Bedürfnisse kann sich das Verhalten der Person situationsabhängig ändern; z.B. wird abhängig von der Befriedigung eines Bedürfnisses unterschiedlich auf dasselbe Objekt reagiert: Für eine bereits gut gesättigte Person stellt z.B. ein großes Stück Sahnetorte nicht unbedingt eine große Versuchung dar, da sie ihren Appetit schon bei der Hauptmahlzeit verausgabt hat. Weil der Hunger, d.h. das Bedürfnis nach Nahrung, bereits gestillt ist, entsteht durch die Torte nicht etwa positiver Affekt, sondern sie kann — nach allzu großem Appetit bei den vorangegangenen Speisen — stattdessen sogar zu einer Quelle negativen Affekts werden. Bei einer am selben Tisch sitzenden Person, die Kuchen besonders gern mag und deshalb zuvor absichtlich wenig gegessen hat, löst das gleiche Kuchenstück jedoch sicherlich positiven Affekt aus. Adäquat an die aktuelle Situation angepasstes Verhalten ergibt sich also durch Kopplung der Befriedigung einzelner Bedürfnisse an die intuitive Verhaltenssteuerung: Durch Abfrage von Bedürfnisbefriedigungen in den *Auslösebedingungen* von Verhaltensweisen wird das Verhalten der Person an die Befriedigung einzelner Bedürfnisse angepasst (siehe Abschnitt 3.).

## 2.4 Wahrgenommene und tatsächliche Befriedigung

Es wird davon ausgegangen, dass Bedürfnisse von ihren Bedürfnismeldern in der Regel “objektiv” darüber informiert werden, ob ein bestimmtes Objekt bedürfnisbefriedigend wirkt. Bedürfnismelder befinden sich allerdings auf einer Ebene, die nicht bewusst wahrgenommen werden kann. Eine Person ist bei ihrer Beurteilung von Objekten auf das Objekterkennungssystem angewiesen, das möglicherweise ungenau arbeitet: Ist die Objekterkennung deaktiviert, werden die Objekte unter Umständen in zu weite Kategorien eingeordnet, so dass wichtige Unterschiede nicht erkannt werden. Dies kann dazu führen, dass Objekte fälschlicherweise als bedürfnisbefriedigend wahrgenommen werden. Umgekehrt wächst bei zu starker Aktivierung des Objekterkennungssystems die Gefahr der *Rigidität*: Kategorien werden zu eng gefasst, da eher unwichtige Unterschiede übertrieben wahrgenommen werden. Dadurch werden bedürfnisbefriedigende Objekte nicht als solche erkannt.

# 3. Die intuitive Verhaltenssteuerung (IVS)

Die Intuitive Verhaltenssteuerung stellt das Repertoire der automatisch verfügbaren Verhaltensroutinen bereit. Dies umfasst einerseits das genetisch vorbereitete Verhalten, andererseits Abläufe, die so intensiv trainiert worden sind, dass sie keine bewusste Kontrolle mehr erfordern. Obwohl automatisches Verhalten ohne Beteiligung bewusster Prozesse ablaufen kann, wird es doch genau an Besonderheiten der aktuellen Situation in Echtzeit angepasst. Für eine hohe Ausführungsgeschwindigkeit ist die gleichzeitige Berücksichtigung einer Vielzahl von Kontextinformationen bei minimalem Zeitbedarf erforderlich. Hierfür eignen sich besonders gut WENN-DANN-Regeln der sogenannten Fuzzy-Logic. Die gewöhnliche zweiwertige Logik wird in der Fuzzy-Logic zusätzlich zu den Wahrheitswerten WAHR und FALSCH um “unscharfe” Wahrheitswerte erweitert. Diese können dazu verwendet werden, um die Merkmale einer Situation sehr viel genauer abzubilden, als dies mit der zweiwertigen Logik möglich ist — man kann gewissermaßen auch die “Halbwahrheiten” zwischen WAHR und FALSCH berücksichtigen. Grundbausteine des Verhaltens stellen die *Basisverhaltensweisen* dar (siehe Abschnitt 3.1). Ein Beispiel für eine “unscharfe” WENN-DANN-Regel wird in Abschnitt 3.2 vorgestellt, um zu

zeigen, wie gut Fuzzy Logic-Regeln zur Charakteristik der Verhaltenssteuerung passen. Die Zusammensetzung von Basisverhaltensweisen zu *Verhaltensprogrammen* folgt in Abschnitt 3.3, in Abschnitt 3.4 wird darauf erläutert, wie durch Verhaltensprogramme eine Verbindung zum noch zu modellierenden kognitiven Makrosystem des *Planens* hergestellt werden kann.

### 3.1 Basisverhaltensweisen

Das Modell der Intuitiven Verhaltenssteuerung stellt automatisches Verhalten durch *Basisverhaltensweisen* und *Verhaltensprogramme* dar. Die Basisverhaltensweisen sind WENN-DANN-Regeln: Bei Erfüllung einer *Bedingung*, die aus mehreren einzelnen Prämissen bestehen kann, wird ein bestimmtes Verhalten ausgeführt (siehe Abb. 3.2): Prämissen werden durch logische

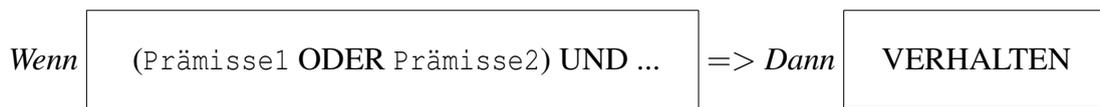
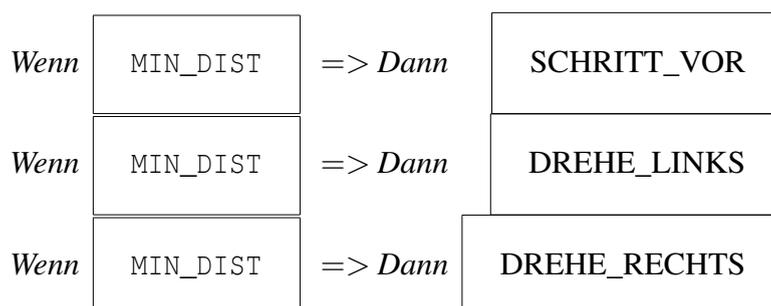


Abbildung 3.2: Basisverhaltensweisen werden durch WENN-DANN-Regeln dargestellt. Die Regeln bestehen aus einer Bedingung und einem Verhalten, das bei Erfüllung der Bedingung ausgeführt wird. Die Bedingung kann aus mehreren Prämissen zusammengesetzt sein, die durch logische Operatoren wie UND, ODER, usw. verknüpft sind. Werden für die Prämissen “unscharfe” Wahrheitswerte zugelassen, kann das Verhalten sehr flexibel an die Erfordernisse der aktuellen Situation angepasst werden.

Operatoren wie UND, ODER, usw. miteinander verknüpft. Wie oben erwähnt sollen die Prämissen statt der Wahrheitswerte WAHR und FALSCH “unscharfe” Wahrheitswerte annehmen können. Hierzu müssen die logischen Operatoren, die zunächst nur für zweiwertige Logik definiert sind, verallgemeinert werden. Da dies im weiteren Verlauf der Arbeit nicht benötigt wird, sei hier auf leicht verständliche Einföhrungsliteratur zur Fuzzy Logic hingewiesen. [18, Traeger]

### 3.2 Beispiel für “unscharfe” WENN-DANN-Regeln

Die Verwendung “unscharfer” Wahrheitswerte soll nun an einem Beispiel erläutert werden: Nehmen wir an, eine Person habe sich vorgenommen ein Ziel in ihrer Umgebung anzusteuern — sicherlich, nachdem die Entscheidung für das Ziel erst einmal gefallen ist, ein Fall für automatisches Verhalten, ohne Beteiligung bewusster Prozesse. Die drei relevanten Basisverhaltensweisen sind dann also



d.h. die Person kann ihre Blickrichtung ändern oder sich einen Schritt nach vorn bewegen. Die drei WENN-DANN-Regeln sind mit nur einer Prämisse versehen: MIN\_DIST ist dann “wahr”, wenn das jeweilige Verhalten die Distanz zum Ziel minimiert. Mit zweiwertiger Logik wäre

---

das Vorgehen klar — man würde feststellen, ob Drehen nach rechts oder links oder Bewegung nach vorn günstiger ist. Dann erhielten alle bis auf eine der drei Prämissen den Wert WAHR, alle anderen dagegen den Wert FALSCH. Es gibt allerdings auch eine Möglichkeit, hier Fuzzy Logic einzusetzen: Hierzu messen wir zunächst die Entfernung  $d^{alt}$  zum Ziel — eine Fähigkeit, die die Intuitive verhaltenssteuerung zur Verfügung hat, da sie eng mit Fähigkeiten zur *räumlichen Orientierung* verbunden ist. Außerdem lassen wir für jede Basisverhaltensweise die Ausführung des Verhaltens “simulieren” und messen jeweils die neue Entfernung  $d_i^{neu}$  zum Zielobjekt. Das simulierte Ausführen von Verhaltensweisen ist ebenfalls typisch für die Intuitive Verhaltenssteuerung: Sie dient zur Abschätzung der kurzfristigen Folgen einzelner automatischer Verhaltensweisen, zwischen denen eine zur Ausführung ausgewählt werden soll. Genau so wird auch in unserem Beispiel verfahren: Wir bilden für jede Basisverhaltensweise den Quotienten

$$a_i = \frac{d^{alt}}{d_i^{neu}},$$

wobei wir  $a_i$  zukünftig als *Aktivierung* der Verhaltensweise  $i$  bezeichnen wollen. Verringert ein Verhalten die Distanz  $d^{alt}$ , so hat  $a_i$  einen Wert größer als 1, ändert sich die Entfernung nicht, erhält die Aktivierung den Wert 1, vergrößert sie sogar die Entfernung zum Zielobjekt, so ergibt sich ein Wert zwischen 0 und 1. Ohne auch nur darüber nachzudenken, welche der verfügbaren Basisverhaltensweisen die Distanz minimiert, erhalten wir mit der Aktivierung also eine Bewertung der einzelner Verhaltensweisen, anhand der wir sowohl ablesen können, ob ein Verhalten die aktuelle Situation verbessert, als auch den Vergleich mit anderen Verhaltensweisen herstellen können: Je größer  $a_i$ , um so besser minimiert das dazugehörige Verhalten die Distanz zum Ziel. Würde die Intuitive Verhaltenssteuerung nun nach der Basisverhaltensweise mit der höchsten Aktivierung suchen... wäre nicht viel gewonnen, denn man hätte auch einfacher herausfinden können, welches Verhalten die Entfernung zum Ziel am stärksten verringert. Worin liegt also nun der Vorteil unseres Vorgehens?

Nehmen wir an, dass unsere Person schon damit zufrieden ist, das Ziel zwar möglichst schnell, aber nicht zwangsläufig in der optimalen Anzahl von Schritten zu erreichen. Dann können die Aktivierungen der “unscharfen” Prämisse dazu verwendet werden, eine Basisverhaltensweise zufällig auszuwählen, wobei die Wahrscheinlichkeit für die Auswahl eines Verhaltens aus der Aktivierung seiner Prämisse bestimmt wird; je höher die Aktivierung, desto höher sollte die Ausführungswahrscheinlichkeit sein. Diese Vorgehensweise garantiert nicht, dass die Prämisse mit der höchsten Aktivierung gewählt wird — man erhält nur mit sehr großer Wahrscheinlichkeit eine Verhaltensweise, die die Distanz halbwegs gut verringert. Dies zahlt sich bei einer großen Anzahl in der Verhaltenssteuerung verfügbarer Verhaltensweisen aus: Dadurch, dass die möglicherweise zeitaufwändige Suche nach einer “optimalen Verhaltensweise” gespart wurde, ist die Intuitive Verhaltenssteuerung in der Lage, Verhalten sehr schnell auszuführen. Dies, aber auch die dabei in Kauf genommene “Großzügigkeit”, nicht auf eine optimale Lösung zuzusteuern, sondern sich mit einer Auswahl halbwegs guter Lösungen zufrieden zu geben, sind Merkmale der Intuitiven Verhaltenssteuerung.

### 3.3 Verhaltensprogramme

Oft ist es nützlich, nicht nur einzelne Basisverhaltensweisen zur Verfügung zu haben, sondern häufig benötigte *Abfolgen mehrerer Basisverhaltensweisen* automatisch ausführen zu können. Dies kann zum Beispiel durch Übung erreicht werden: Beim Autofahren werden einfache Bewegungsabläufe wie das Treten verschiedener Pedale, die Bedienung des Schalthebels und die

---

Führung des Lenkrades automatisch ausgeführt — nachdem man dies in der Fahrschule recht lange geübt hat. Dass die Bewegungsabläufe zunächst nicht automatisch ablaufen, erkennt man sehr deutlich, wenn man einen Fahrschüler beobachtet, der immer wieder versucht, zu schalten, *bevor* er die Kupplung getreten hat. Einem geübten Autofahrer passiert das nie — er *kann* den Fehler vermutlich kaum mehr machen, weil er den Schaltvorgang ohne bewusste Kontrolle mit dem Treten der Kupplung beginnt. Solche Sequenzen von Basisverhaltensweisen werden im Modell durch *Verhaltensprogramme* dargestellt (siehe Abb. 3.3): Ein Verhaltensprogramm führt

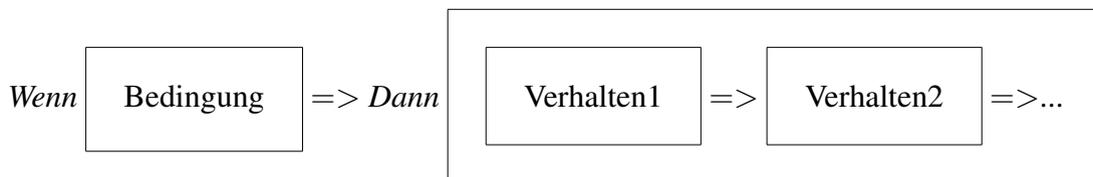


Abbildung 3.3: Verhaltensprogramme bestehen aus einer Abfolge mehrerer Basisverhaltensweisen. Bei erfüllter Bedingung werden nacheinander die einzelnen Basisverhaltensweisen ausgeführt, ohne, dass die Bedingungen der einzelnen Basisverhaltensweisen erfüllt sein müssen.

bei Erfüllung der Bedingung gleich eine ganze Folge von Basisverhaltensweisen aus, ohne, dass die einzelnen Bedingungen der zum Programm gehörigen Verhaltensweisen geprüft werden müssen. Durch Verhaltensprogramme werden also Basisverhaltensweisen — üblicherweise unter schwerer zu erfüllenden Auslösungsbedingungen als für das Ausführen einer einzelnen Basisverhaltensweise — zu größeren Einheiten zusammengefasst. Das Verhalten wirkt dadurch von außen betrachtet “intelligenter”, obwohl es nichts von seiner Charakteristik automatischen Verhaltens verliert.

### 3.4 Verhaltensprogramme und Planen

Dass Verhaltensprogramme aus Sequenzen einfacherer Basisverhaltensweisen bestehen, bereitet schon auf eine Erweiterung des Modells um das Planen vor. Das Planen wird für das Umsetzen schwieriger Absichten benötigt, wobei Probleme auftauchen können, die nicht allein durch automatische Verhaltensroutinen überwunden werden können. Planen dient — wie der Name dieses kognitiven Systems sagt — dazu, Pläne zu schmieden; also nach Schritt für Schritt realisierbaren Handlungsabläufen zu suchen, die letztendlich im einzelnen durch das Ausführen von verfügbaren automatischen Verhaltensroutinen implementiert werden müssen. Mit Sequenzen von Basisverhaltensweisen besteht also schon eine “gemeinsame Sprache” von Planen und Intuitiver Verhaltenssteuerung: Pläne können einerseits in Form von Verhaltensprogrammen an die Verhaltenssteuerung übertragen werden. Zum Entwickeln von Plänen muss das Planen jedoch andererseits das verfügbare Verhalten in einer dem Problem angemessenen Form erhalten. Müsste ein Autofahrer das Umfahren eines Staus anhand der Basisverhaltensweisen “Betätigen des Kupplungspedals”, “Gasgeben”, “Bremsen”, “Schalten” und “Lenken” planen, würde er vermutlich mindestens bis zur Auflösung des Staus mit Planen beschäftigt sein. Beim Arbeiten mit Verhaltensprogrammen wie “Geradeaus fahren”, “Links abbiegen” und “Rechts abbiegen”, die jeder Autofahrer beherrscht, erscheint das Finden einer Umgehung schon weit weniger aussichtslos.

## 4. Affektdynamik

Nach der Vorstellung der Modelle der Objekterkennung und des Verhaltens, muss noch geklärt werden, in welchen Situationen welches kognitive System aktiv wird. Situationen werden durch Affekt bewertet, deshalb liegt es nahe, Wahrnehmung und Verhalten gemäß der Affektlage auszurichten. Genau dies tut Kuhl mit der Formulierung seiner *Modulationsannahmen* [9, Kuhl, Kapitel 5], die die Aktivität von Objekterkennung, Fühlen, Verhalten und Planen in Abhängigkeit von positivem und des negativem Affekts postulieren.

Ein von Bröcker und Kuhl entwickeltes Differentialgleichungsmodell (siehe [2, Bröcker]) stellt die *Affektdynamik*, die sich durch die Modulationsannahmen ergibt, bereits gut dar. In [13] wurde deshalb nach dem Vorbild des Differentialgleichungsmodells mit der Entwicklung eines sich äquivalent verhaltenden Differenzgleichungsmodells begonnen. Das Modell soll die gemäß des STAR-Modells (siehe Abb. 4.4) zu erwartende Dominanz einzelner kognitiver Systeme in Abhängigkeit von den sogenannten *Affektsensibilitäten* beschreiben. Affektsensibilitäten be-

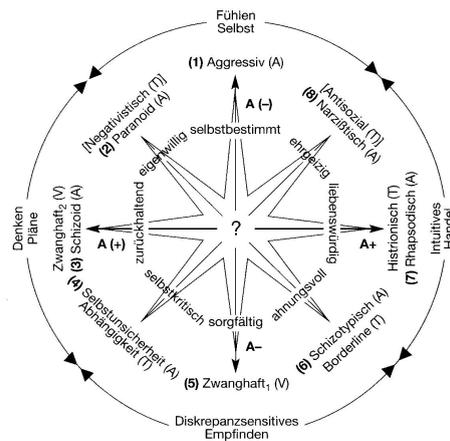


Abbildung 4.4: Das STAR-Modell aus [9] ordnet Affektsensibilitäten dominante psychische Makrosysteme zu und leitet aus den charakteristischen Eigenschaften der dominierenden Makrosysteme Persönlichkeitsstile ab.

zeichnen die Tendenz, Affekt, der oberhalb des Ruheniveaus 1 ist, schnell oder langsam wieder auf das Ruheniveau herabzuregulieren. Ist eine Person sensibel für positiven Affekt (A+), bleibt erhöhter positiver Affekt vergleichsweise lange erhöht (siehe Abb. 4.5), wird positiver Affekt gedämpft, wird er schnell wieder auf das Ruheniveau zurückgeführt (siehe Abb. 4.6).

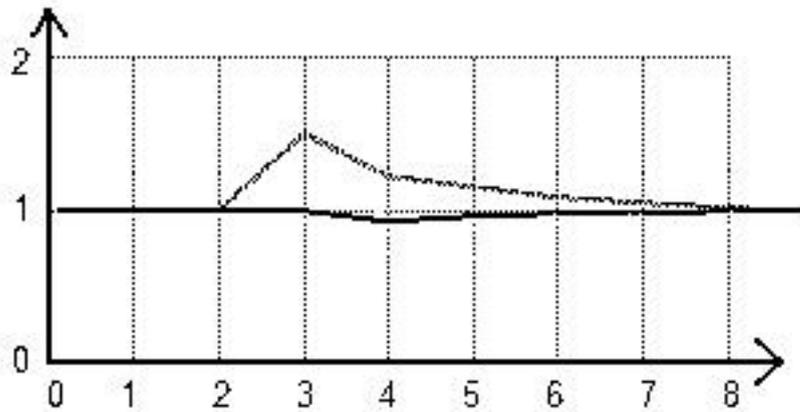


Abbildung 4.5: Die Abbildung zeigt die Reaktion auf eine Affektstörung bei Sensibilität für diesen Affekt: Der Affekt bleibt einige Zeitschritte über dem Ruheniveau 1.

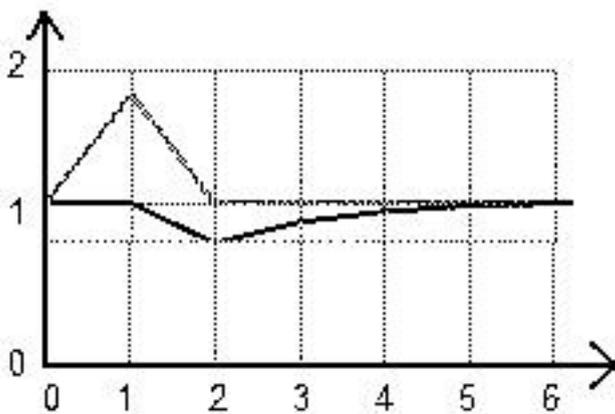


Abbildung 4.6: Die Abbildung zeigt die Reaktion auf eine Affektstörung, falls die Person nicht sensibel für den Affekt ist: Der Affekt bleibt nur einen Zeitschritt über dem Ruheniveau 1.

---

## 4.1 Differenzgleichungsmodell der Affektdynamik

Das Differenzgleichungssystem (4.6) - (4.9)

$$F_{n+1} = \frac{E_n}{A_-} F_n \quad (4.6)$$

$$E_{n+1} = E_n + 0.5 \left( A_- - \frac{E_n^2}{A_-} \right) F_n^2 \quad (4.7)$$

$$P_{n+1} = \frac{V_n}{A_+} P_n \quad (4.8)$$

$$V_{n+1} = V_n + 0.5 \left( A_+ - \frac{V_n^2}{A_+} \right) P_n^2 \quad (4.9)$$

wurde in [13] und [14] entwickelt, um das Verhalten des von Bröcker und Kuhl entwickelten Differentialgleichungsmodell in ein zeitdiskretes Modell zu übertragen. Die Gleichungen (4.8) und (4.9) stehen mit den Gleichungen (4.6) und (4.7) zunächst nicht in Beziehung. Die Verbindung wird durch zwei Gleichungen für den Affekt hergestellt:

$$b_+ = \begin{cases} \frac{2}{1+e^{-s_{++}(F-P)}} & F - P \geq 0 \\ \frac{2}{1+e^{-s_{+-}(F-P)}} & F - P < 0 \end{cases} \quad (4.10)$$

$$b_- = \begin{cases} \frac{2}{1+e^{-s_{-+}(F-P)}} & F - P < 0 \\ \frac{2}{1+e^{-s_{--}(F-P)}} & F - P \geq 0 \end{cases} \quad (4.11)$$

Hierbei drücken  $s_{++}$ ,  $s_{+-}$ ,  $s_{-+}$  und  $s_{--}$  die Affektsensibilitäten der Person aus:  $s_{++}$  und  $s_{+-}$  beschreiben die Tendenz, positiven Affekt zu erhalten bzw. zu dämpfen,  $s_{-+}$  und  $s_{--}$  haben die analoge Bedeutung für negativen Affekt.

## 4.2 Korrekturen des Differenzgleichungsmodells

Problematisch am vorliegenden Modell ist die Instabilität der Gleichungen (4.6) und (4.8). Dies führt zu unerwünschtem Verhalten des Modells, wenn es über viele Zeitschritte läuft: Das erwünschte Stabilitätsniveau der Aktivierungen von Fühlen und Planen verschiebt sich zu immer höheren Werten. Dadurch wird schließlich das gesamte Differenzgleichungssystem aus dem Gleichgewicht gebracht.

Die Lösung des Problems besteht darin, die gleichzeitige Aktivierung der komplementären Systeme Fühlen und Objekterkennung sowie Planen und Verhaltenssteuerung zu verhindern: Erreichen zwei komplementäre Systeme gleichzeitig einen Wert oberhalb oder unterhalb von 1, werden mithilfe einer Fallunterscheidung, die auf den Annahmen der PSI-Theorie beruht, die Werte wie folgt behandelt: Falls die Aktivierungen von Fühlen und Planen gleichzeitig den Wert 1 überschreiten, wird kontrolliert, welcher Wert nach der aktuellen Affektlage am ehesten "überhört" ist:

$$F_{n+1}^* = \begin{cases} F_{n+1}, & \text{falls } A_+ > A_- \\ 1, & \text{falls } A_+ \leq A_- \end{cases} \quad (4.12)$$

$$P_{n+1}^* = \begin{cases} P_{n+1}, & \text{falls } A_+ < A_- \\ 1, & \text{falls } A_+ \geq A_- \end{cases} \quad (4.13)$$

Da Fühlen durch negativen Affekt gehemmt wird, wird es auf 1 gesetzt, falls negativer Affekt überwiegt und sowohl die Aktivierung des Fühlens als auch des Planens über dem Niveau 1 liegen. Analog wird der Wert des Planens auf 1 herabgesetzt, wenn Fühlen und Planen Aktivierungen oberhalb von 1 erreicht haben und positiver Affekt überwiegt. Die korrigierten Aktivierungen werden mit  $F_{n+1}^*$  und  $P_{n+1}^*$  bezeichnet, die nach (4.6)-(4.9) berechneten dagegen mit  $F_{n+1}$  und  $P_{n+1}$ .

Analog werden die Aktivierungen von Verhaltenssteuerung  $V$  und Objekterkennung  $E$  behandelt:

$$V_{n+1}^* = \begin{cases} V_{n+1}, & \text{falls } A+ > A- \\ 1, & \text{falls } A+ \leq A- \end{cases} \quad (4.14)$$

$$E_{n+1}^* = \begin{cases} E_{n+1}, & \text{falls } A+ < A- \\ 1, & \text{falls } A+ \geq A- \end{cases} \quad (4.15)$$

Die Korrekturen für den Fall, dass zwei Systeme unter das Niveau 1 reguliert werden, können leicht hergeleitet werden; hier ist zu untersuchen, welche Systeme nach der aktuellen Affektlage eine "zu niedrige" Aktivierung aufweisen und deshalb auf das Ruhenniveau 1 heraufgesetzt werden sollten.

### 4.3 Bedürfnisgenerierter und antizipierter Affekt

Setzt man alle Gleichungen (4.6)-(4.9) auf den Anfangswert 1, so verändern sich die Aktivierungen nicht. Hierzu ist eine Störung durch Veränderung einer Affektsorte nötig. Unter welchen Bedingungen sollte Affekt gestört werden, d.h. auf ein höheres Niveau gesetzt werden? Im Modell sind zwei Quellen für Affekt vorgesehen: Einerseits Affekt, der anhand der Befriedigungslage der Bedürfnisse (siehe Gleichung (2.3)) berechnet wird — dies wird hier mit *bedürfnisgeneriertem Affekt* bezeichnet. In den durchgeführten Anwendungen des Modells wurde diese Möglichkeit der Affektentstehung bisher jedoch selten verwendet. Die andere Möglichkeit der Affektgenerierung ist der *antizipierte Affekt*, der durch Objekte der Umgebung zustandekommt. Hierzu wird für ein Objekt, auf das sich zum aktuellen Zeitschritt die Aufmerksamkeit richtet, der durch die Bedürfnisse mit diesem Objekt verknüpfte Affekt berechnet (siehe Abschnitt 2.1). Dieser Affekt kann als Antizipation des zu erwartenden Affekts gesehen werden, der dann ausgelöst würde, wenn das Objekt die "Chance" zur Bedürfnisbefriedigung bekäme. Antizipierter positiver Affekt würde beispielsweise auftreten, wenn das Objekterkennungssystem dem Bedürfnis nach Nahrung das Auftauchen eines essbaren Objektes meldet.

### 4.4 Parametrisierung

Wie das Differentialgleichungsmodell von Bröcker und Kuhl kann auch mit dem hier entwickelten Differenzgleichungsmodell die Affektdynamik der Persönlichkeitstypen des STAR-Modells simuliert werden. Hierzu sind die Parameter gemäß folgender Tabelle zu wählen: Die Parametrisierung entspricht den in Abb. 4.4 bildlich dargestellten Ausprägungen der Affektsensibilitäten. Aus technischen Gründen muss bei Heraufregulierung einer Affektsorte auch ein niedriger Wert für die Herabregulierung dieser Affektsorte gewählt werden; ansonsten würde dieser Affekt bei Auftreten der anderen Affektsorte nicht herabreguliert, sondern auf dem Ruhenniveau 1 verbleiben.

---

**Tabelle 2.1: Parametrisierung des Differenzgleichungsmodells der Affektdynamik für die Persönlichkeitsstile des STAR-Modells**

Persönlichkeitsstil	posAffUp	negAffUp	posAffDn	ngAffDn
selbstbestimmt <sub>1</sub>	0.3	0	0.2	0.8
eigenwillig	0	0	0.6	0.6
zurückhaltend	0	0.3	0.8	0.2
selbstkritisch	0	1	0.8	0.2
sorgfältig	0	1	0.3	0.2
ahnungsvoll	1	1	0.3	0.3
optimistisch	1	0	0.2	0.3
ehrgeizig	1	0	0.2	0.8

# Kapitel 3

## Ein Modell für Extensionsgedächtnis, Fühlen und Selbst

Dieser Abschnitt beschreibt detailliert die Modellierung des Fühlens ausgehend von den theoretischen Grundlagen (Abschnitt 1.). Daran schließt sich ein Abschnitt über neuronale Netzwerkmodelle an, die auf ihre Eignung zur Modellierung der zuvor beschriebenen Eigenschaften untersucht werden (Abschnitt 2.). Die Verknüpfung der in diesem Abschnitt vorbereiteten Komponenten zu einem Modell des Fühlens wird in Abschnitt 3. erläutert. Den Abschluss des Kapitels bildet Durchführung und Diskussion eines Simulationsexperiments (Abschnitt 4.).

### 1. Fühlen, Selbst und Extensionsgedächtnis in der PSI-Theorie

#### 1.1 Das Fühlen

Ein Computermodell eines so subjektiven Phänomens wie des Fühlens scheint zunächst aussichtslos zu sein. Versteht man Fühlen als die “Gefühlswelt” einer Persönlichkeit, ist ein wissenschaftlicher Zugang sehr schwierig, da sich hier nur das Fühlen einzelner Persönlichkeiten beschreiben ließe. Allgemeine Aussagen wären nur über den Vergleich des Fühlens vieler Persönlichkeiten möglich. Viel einfacher ist der Zugang zum Fühlen allerdings, wenn man es als kognitive Funktion im Sinne Jungs versteht. Fühlen ist in dieser Betrachtungsweise eine von vier möglichen Arten, Informationen der Außen- und Innenwelt einer Persönlichkeit zu verarbeiten (s. [8, Jung]). Dieser Gedanke wird in der PSI-Theorie aufgegriffen, jedoch erhalten die Begriffe *Empfinden*, *Intuieren*, *Denken* und eben das *Fühlen* eine von den Vorstellungen Jungs teilweise mehr oder weniger stark abweichende Bedeutung. Im Gegensatz zu den elementaren Systemen der Objekterkennung und der Verhaltenssteuerung, die Empfinden und Intuieren entsprechen sollen, ist das Fühlen eine höher entwickelte Form der Informationsverarbeitung: Eine große Menge von der Objekterkennung isoliert wahrgenommener Einzelinformationen wird durch das Fühlen in ein weitläufiges assoziatives Netzwerk eingebettet. Hierbei werden Repräsentationen des eigenen Zustands, d.h. zum Beispiel der Status der Bedürfnisse oder die Affekt- und Bedürfnislage in bestimmten Situationen mit Verhaltensweisen und den aus ihrer Ausführung resultierenden Handlungsergebnissen verknüpft.<sup>1</sup> Mit Fühlen in diesem Sinne ist das “Etwas-im-Gefühl-haben”, also der Zugriff auf einen großen Vorrat autobiographischen

---

<sup>1</sup>“Mit Fühlen ist eine hochinferente Form der Wissensrepräsentation bezeichnet, die [...] durch *Konfigurationen* von Erfahrungen, wie z.B. umfassende Erlebnisepisoden oder ausgedehnte Netzwerke von Selbstaspekten oder Handlungsmöglichkeiten (kognitiv-emotionale Landschaften) [charakterisiert ist]. Fühlen abstrahiert einer-

---

Erfahrungswissens gemeint, das durch Abstraktion aus vielen einzelnen Wahrnehmungen des Objekterkennungssystems gewonnen worden ist. Dies erlaubt es, die kognitive Funktion des Fühlens von den *Gefühlen* zu trennen, die subjektiv stark unterschiedlich sein können. Gefühle ergeben sich in diesem theoretischen Konzept durch affektive Bewertung der Inhalte des Fühlens. Mit den bereits vorhandenen Modellen zu Affekt, Bedürfnissen und Objekten (siehe Kapitel 2, [15, Siekmann]), erscheint die Aufgabe, ein Computermodell zu einem auf diese Weise verstandenen Fühlen zu entwickeln, bereits weit weniger problematisch.

## 1.2 Das Selbst

Eine noch größere Herausforderung ist die Modellierung des sogenannten Selbst-Systems. Das Selbst wird in der PSI-Theorie als der Bestandteil einer Persönlichkeit angesehen, der unseren Charakter am entscheidendsten ausmacht. Es zeigt sich erneut, dass wie beim Fühlen nicht die Inhalte einzelner psychischer Systeme Hauptgegenstand der PSI-Theorie sind, sondern dass auch das Selbst eine psychische Funktion darstellt, deren Stärke im Vergleich zu und deren Zusammenwirken mit anderen psychischen Systemen untersucht wird.<sup>2</sup> Letztendlich entscheidet die Intaktheit des Selbst darüber, ob eine Person durch Fixierung auf bestimmte Systemkonstellationen eine Persönlichkeitsstörung entwickelt oder “gesund” ist.<sup>3</sup>

### Funktion des Selbst

Das Selbst stellt “die höchste Stufe der Integration der für die Befriedigung sämtlicher Bedürfnisse und Anliegen wichtigen Informationen” [9, Kuhl, S. 716] dar. Diese Repräsentation eigener Zustände wird in der PSI-Theorie mit dem Begriff der *integrierten Selbstrepräsentation* (siehe Abschnitt 1.4) bezeichnet. Im Persönlichkeitsmodell der PSI-Theorie übernimmt das Selbst die Rolle einer zentralen Exekutive, die Eingriffsmöglichkeiten auf allen Systemebenen hat. Seine Aufgabe ist es, “andere Systeme so zu koordinieren, daß immer wieder Handlungsmöglichkeiten gefunden werden, die die *simultane* Befriedigung möglichst vieler Bedürfnisse und Anliegen im Rahmen der Möglichkeiten ausschöpft”. [9, Kuhl, S. 716]

### Allgemeine Ziele

Hieraus ergibt sich, dass das Selbst *allgemeine Ziele* unterstützt, die im Gegensatz zu expliziten Zielen (siehe [9, Kuhl, S.158]) nicht auf die Befriedigung eines einzelnen Bedürfnisses oder das Erreichen eines fest vorgegebenen Handlungsergebnisses ausgerichtet sind. Kuhl definiert

---

seits Gemeinsamkeiten aus *episodischen Erfahrungen* [...], ohne sie ihrer emotionalen Komponenten zu berauben, und integriert andererseits immer neue Empfindungen und Einzelobjekte, (Kuhl, S. 626) (die bislang noch nicht zu dem Horizont der im Extensionsgedächtnis gespeicherten Erwartungen, Wünsche, Bedürfnisse und weiterer Selbstaspekte gehörten.)” [9, Kuhl, S.626f.]

<sup>2</sup>“*Persönlichkeit* [bezeichnet] die dispositionellen Stärkeverhältnisse und die daraus resultierenden Interaktionsmuster einzelner psychischer Systeme [...], während *Charakter* den geformten Inhalt einzelner psychischer Systeme bezeichnet, besonders den Inhalt des integrierten Selbst mit seinen persönlichen Werten und Einstellungen.” [9, Kuhl, S. 93]

<sup>3</sup>Eine theoretische Annahme der PSI-Theorie ist, “daß die *Hemmung von Selbststeuerungsfunktionen* ein krankheitsübergreifendes Symptom einer psychischen Dekompensation ist (analog zum Fieber in der Medizin als unspezifischem Indikator für die Aktivität des Abwehrsystems).” [9, Kuhl, S.996] Im Gegensatz dazu kann “Bei intakten Selbststeuerungsfunktionen [...] selbst bei Menschen, die beträchtliche affektive oder kognitive Einseitigkeiten entwickelt haben, sogar eine besonders tiefgreifende und harmonische Entwicklung der Persönlichkeit *begünstigt werden*.” [9, Kuhl, S.991]

---

ein allgemeines Ziel als “ganzheitliche Repräsentation eines Netzwerks potentiell akzeptabler Handlungsergebnisse und gleichwertiger Alternativziele” [9, Kuhl, S.151]. Der Verknüpfungsgrad solcher Netzwerke, die Teil des sogenannten Extensionsgedächtnisses sind, auf das später in Abschnitt 1.4 eingegangen werden wird, ist so hoch, dass ein allgemeines Ziel im Gegensatz zu einem expliziten Ziel vom Bewusstsein weder vollständig wahrgenommen noch sprachlich repräsentiert werden kann.<sup>4</sup>

### Allgemeine Ziele sind eher “selbstgewollt” als explizite Ziele

Allgemeine Ziele sind in ein Netzwerk aus Selbstrepräsentationen eingebettet. Deshalb besteht weniger als bei bewussten, sprachlich repräsentierten, “expliziten” Zielen die Gefahr, dass sie den Bedürfnissen, Werten und Einstellungen einer Persönlichkeit widersprechen.<sup>5</sup> Dies ist um so bedeutsamer, da die für das Verfolgen expliziter Ziele verantwortliche psychische Funktion des Denkens möglicherweise die Wahrnehmung unterdrückt, dass ein explizites Ziel im Widerspruch zum Selbst steht. Beim Verfolgen eines expliziten Ziels werden Anteile des Selbst unterdrückt, die der Verwirklichung des Ziels entgegenstehen — die Person handelt “selbstdiszipliniert”. So kann die paradoxe Situation vorkommen, dass eine Persönlichkeit ein explizites Ziel intentional verfolgt, ohne es eigentlich zu “wollen”.<sup>6</sup> Damit ist die Untersuchung allgemeiner Ziele besonders wichtig, da deren Verwirklichung im allgemeinen eher dem “Wohl” der Person dienen, als explizite Ziele.

### Selbstmotivierung und Selbstberuhigung

Für die Umsetzung von schwierigen Zielen, die nicht durch das Verhaltensrepertoire der Intuitiven Verhaltenssteuerung erreicht werden können (siehe Abschnitt 3.), ist oft die Bereitstellung zusätzlichen positiven Affekts oder die Herabregulierung negativen Affekts erforderlich.<sup>7</sup> Nun gibt es durchaus auch andere Möglichkeiten, negativem Affekt oder der Hemmung von positivem Affekt “aus dem Weg zu gehen”. Dies ist ausführlich in [9, Kuhl, Kapitel 11] ausgearbeitet. Hier sei nur die Strategie des *positiven Ablenkens* dargestellt, um den Unterschied zu Selbstmotivierung und Selbstberuhigung zu illustrieren: Beim *positiven Ablenken* findet eine “Flucht in

---

<sup>4</sup>“Auf der Ebene des bewußten Erlebens sollte sich der intakte Zugriff auf das Extensionsgedächtnis besonders in *positiven Kontrollerwartungen* niederschlagen: Wer den Zugriff auf dieses assoziative Netzwerk hat, das viele Handlungsmöglichkeiten zur Verwirklichung eines allgemeinen Ziels bereithält, kann zwar nicht alle diese Handlungsmöglichkeiten aufzählen (da es sich um ein implizites Gedächtnissystem handelt), trotzdem aber sagen, daß er einfach “spürt”, daß er das Ziel “irgendwie” erreichen kann.” [9, Kuhl, S. 155]

<sup>5</sup>“Wie aus der Diskussion der *Selbstkontrolle* hervorgeht, wäre es falsch, aus diesem Begriff zu schließen, daß hier die Verfolgung eigener, d.h. auf der Ebene des Selbst repräsentierter Bedürfnisse im Vordergrund steht. Im Gegenteil: Das “Selbst” ist bei der “Selbst-Kontrolle” Gegenstand oder “Opfer” der Kontrolle, nicht Subjekt oder Urheber der Kontrolle. Das kann immer dann adaptiv sein, wenn es um die Verfolgung von Zielen mit hoher Priorität geht. Dann kann es sich — zumindest langfristig — auszahlen, das Selbstsystem mit dem von ihm geleisteten Überblick über sämtliche innere Regungen, Gefühle, Interessen und Bedürfnisse vorübergehend außer Kraft zu setzen, um “Ablenkungen” zu vermeiden.” [9, Kuhl, S.713]

<sup>6</sup>“Wird das Selbstsystem andauernd unterdrückt, können Ziele und Handlungsabsichtigen gar nicht auf *Selbstkompatibilität* geprüft und ggfs. ins Selbst integriert werden, ja es kann sogar zu *fehlinformierter Introjektion* kommen, die darin besteht, daß die Person fälschlicherweise ihre Ziele auf der Ich-Ebene für selbstintegriert hält.” [9, Kuhl, S.134], zitiert (Kuhl & Kazèn, 1994)

<sup>7</sup>“[...] wenn bei einer relativ unangenehmen Aufgabe zu wenig positive Anreizmotivation da ist, werden negative Affekte aus dem Bestrafungssystem durch Generieren positiven Affekts “beschwichtigt” oder sogar verborgene Anreizmomente im Belohnungssystem aktiviert (*Motivationsregulierung* oder: *Selbstmotivierung*).” [9, Kuhl, S.698]

---

den positiven Affekt” statt, d.h. es wird bei Auftreten negativen Affekts oder der Hemmung positiven Affekts *automatisch* mit der Generierung positiven Affekts reagiert. Diese Art der Problembewältigung erscheint sehr einfach, allerdings besteht hier die Gefahr, dass sie gerade in wirklich schwierigen Situationen scheitert: “Beschönigen” einer Situation oder “Bagatellisieren” von Schwierigkeiten funktioniert sehr gut, wenn die Probleme nicht besonders groß sind, versagt jedoch auf mittlerem und erst recht auf hohem Stressniveau völlig. (siehe [9, Kuhl, S.424-426]) Der *selbstgesteuerten Auseinandersetzung* gesteht Kuhl dagegen zu, “daß sie die einzige *inhaltlich* echte Bewältigungsform ist, da Erfahrungswissen eingebracht wird, daß den bedrohlichen Charakter der angstauslösenden Wahrnehmungen entschärft, entweder durch das Auffinden einer passenden Handlungsmöglichkeit oder durch Umbewertung und Sinnstiftung (z.B. bei leidvollen Erfahrungen, die sich nicht mehr rückgängig machen lassen).” [9, Kuhl, S.737] Allerdings setzt die Motivationsregulierung mithilfe des Selbst-Systems die Konfrontation mit negativen Affekt voraus — das Selbst muss erkennen, dass sein Eingreifen notwendig ist —, so dass diese Bewältigungsform dem *positiven Ablenken* in einer Phase niedriger Stressintensität (scheinbar) unterlegen ist.

### 1.3 Kongruenzorientierte Aufmerksamkeit (Vigilanz)

Allgemeine Ziele sind nicht nur mit dem Selbst, sondern auch mit der kognitiven Funktion des Fühlens eng verknüpft. Die *kongruenzorientierte Aufmerksamkeit* oder *Vigilanz* versucht “Kongruenz” mit dem Selbst, d.h. den Bedürfnissen, Einstellungen und Werten einer Persönlichkeit herzustellen. Ihre Funktion besteht darin, selbst auf den ersten Blick nicht gleich erkennbare “Kongruenzen” der aktuellen Lage mit Situationsmerkmalen zu finden, die das Verfolgen eines allgemeinen Ziels ermöglichen. Diese Art von Aufmerksamkeit entspricht dem “Warten auf seltene Gelegenheiten”, die vor allem durch in einem langjährigen Prozess erworbenes Erfahrungswissen erkannt werden können. Im Gegensatz zur *absichtsorientierten Aufmerksamkeit*, dem bewussten Verfolgen eines fest vorgegebenen, “expliziten” Ziels, ist bei kongruenzorientierter Aufmerksamkeit oft nicht klar, welche Handlungsschritte zum Erreichen eines möglicherweise noch nicht einmal genau definierten Ziels ausgeführt werden müssen.<sup>8</sup> Das Erkennen von Gelegenheiten für das Erreichen allgemeiner Ziele und die Steuerung des hierfür erforderlichen, fein an die vorliegende Situation angepassten Verhaltens, ist also an eine gute Entwicklung der Fähigkeiten des Fühlens geknüpft.

### 1.4 Das Extensionsgedächtnis (EG)

Fühlen und Selbst sind eng verbunden mit dem sogenannten Extensionsgedächtnis. Dieses stellt “ein so breites assoziatives Netzwerk impliziten Wissens bereit, daß eine vollständige Übersetzung in die propositionale Sprache des sequentiellen Denkens gar nicht möglich ist” [9, Kuhl, S.660]. Die Unterscheidung von implizitem und explizitem Gedächtnis stammt von dem amerikanischen Psychologen Daniel L. Schacter. Er bezeichnet als *explizit* eine Gedächtnisform, die auch *deklaratives Gedächtnis* genannt wird. Diese Art des Gedächtnisses umfasst eine Sammlung von Fakten, die bewusst erinnert werden können. Im *impliziten Gedächtnis* sind dagegen

---

<sup>8</sup>“[...]das System “Fühlen” [ist] in der PSI-Theorie eng mit einer Form von Aufmerksamkeit vernetzt, die ich “kongruenzorientiert” nenne [...] Motivationspsychologisch ist diese Aufmerksamkeitsfunktion besonders für die Umsetzung von Zielen relevant, für die keine konkreten Handlungsschritte oder Auslösebedingungen festgelegt sind.[...] Sie [erlaubt] Handlungsantizipation und Zielbildung auch dann [...], wenn es noch gar nicht möglich oder sinnvoll ist, konkrete Ausführungsbedingungen festzulegen.” [9, Kuhl, S.667f.]

---

Erfahrungen repräsentiert, die auch ohne bewussten Zugang zur Verfügung stehen.<sup>9</sup>

### **Integrierte Selbstrepräsentationen**

Dadurch, dass autobiographische Episoden Bestandteil dieses assoziativen Netzwerks sind, ist das Selbst teilweise im Extensionsgedächtnis repräsentiert<sup>10</sup>, (s. auch [9, Kuhl, S.719ff]). Die Verbindung des Selbst mit dem Extensionsgedächtnis besteht in den *integrierten Selbstrepräsentationen*. Diese zeichnen sich aus durch “eine hohe Stufe der *Integration* von Einzelempfindungen über eigene Bedürfnisse, Gefühle, Körperwahrnehmungen, Präferenzen, Werte und den Erfahrungskontext [...], in dem solche Erlebnisse gemacht wurden”[9, Kuhl, S. 151].

### **Die Bildung integrierter Selbstrepräsentationen**

Demgegenüber ist das Bilden und die Erweiterung bestehender Selbstrepräsentationen durch Verarbeitung isolierter Einzelempfindungen der Objekterkennung eine der Aufgaben des Fühlens.<sup>11</sup> Das Fühlen stellt dabei mithilfe des impliziten Erfahrungswissens des Extensionsgedächtnisses möglicherweise auch sehr weit hergeholte Zusammenhänge zwischen Wahrnehmungseindrücken her, die durch die Informationsverarbeitung der Objekterkennung nicht zutage treten können, ja sogar unterdrückt werden. Dies wird in Abschnitt 1.5 weiter ausgeführt. Die letzten beiden Abschnitte zeigten, dass das Extensionsgedächtnis sozusagen die “Hardware” für das Fühlen und einen Teils des Selbst (das ganzheitliche, “implizite” Selbst, siehe [9, Kuhl, Kapitel 14]) bereitstellt. Ein Computermodell des Fühlens und des Selbst muss also mit der Implementierung des Extensionsgedächtnisses beginnen. Fühlen und Selbst ergeben sich dann durch Festlegung des Zugriffs auf dieses simulierte Gedächtnis.

## **1.5 Fühlen als Zusammenspiel von OES und EG**

Bevor neuronale Netzwerkmodelle auf ihre Eignung zur Simulation des Extensionsgedächtnis geprüft werden, muss noch der Informationsaustausch zwischen Objekterkennung und Extensionsgedächtnis genauer betrachtet werden. Wie bereits im Abschnitt über das Fühlen gesagt, besteht das Extensionsgedächtnis aus einer assoziativen Verknüpfung einer Vielzahl von in einem langen Prozess gesammelten Einzeleindrücken des Objekterkennungssystems. Kuhl greift in seiner Darstellung der Informationsübertragung vom Objekterkennungssystem zum Extensionsgedächtnis stark auf einen einflussreichen Artikel von James McClelland, Bruce McNaughton und Randall O’Reilly zurück, die in [10, McClelland et al.] eine Theorie der Rolle des

---

<sup>9</sup>“Explicit memory involves the conscious, intentional recollection of previous experiences, what we tend to think of as memory in our everyday lives. It may involve reliving or reexperiencing past events. Implicit memory refers to nonconscious, unintentional influences of past experiences on current behavior and performance.” [12, Schacter]

<sup>10</sup>“Das Extensionsgedächtnis integriert ausgedehnte Konfigurationen von Einzelobjekten und selbstrelevanten autobiographischen Episoden zu ganzheitlichen “Erlebnislandschaften”, einschließlich der in diesen Episoden erlebten Emotionen und Körperempfindungen und ist deshalb eine gute Basis für die Steuerung von Emotionen (Emotionsbewältigung und Selbstmotivierung).” [9, Kuhl, S.162]

<sup>11</sup>“Mit dem Begriff des Fühlens sind Strukturen gemeint, die durch die Interaktion des Extensionsgedächtnisses mit diskrepanten Objektwahrnehmungen in der Vergangenheit gewachsen sind[...]. Der Begriff des Extensionsgedächtnisses kann dagegen unabhängig davon verwendet werden, wie viel Kommunikation mit diskrepanten Objektwahrnehmungen in der Vergangenheit stattgefunden hat: Wurden wenig Unstimmigkeiten (d.h. diskrepante “Objekte”) integriert, so sind die entsprechenden Strukturen des EG entsprechend flach.” [9, Kuhl, S.627, s.S.127]

---

Hippocampus bei der Übertragung von Gedächtnisinhalten in den Neocortex entwickeln. Der Hippocampus nimmt — wie sein natürliches Vorbild — auch in der PSI-Theorie die Rolle eines Pufferspeichers ein: Er reguliert den Informationsfluss zwischen Objekterkennungssystem und Extensionsgedächtnis. Mit einer gewissen Vorsicht wird hier wie auch an anderen Stellen der PSI-Theorie versucht, bestimmte Systeme der Theorie neurophysiologisch zu lokalisieren; in diesem Fall das Extensionsgedächtnis im Neocortex, während die Rolle des in [9, Kuhl, Kapitel 11] beschriebenen Umschalters zwischen automatischen und höheren Prozessen weitgehend dem Hippocampus zufällt. Da der erwähnte Artikel für die Theorie des Fühlens in [9, Kuhl, Kapitel 11 & 13] eine zentrale Rolle einnimmt und er im hier vorgestellten Computermodell zentrale Berücksichtigung findet, soll er hier zusammenfassend dargestellt werden.

### **Der Hippocampus als “Lehrer” des Langzeitgedächtnisses**

Die Autoren des Artikels gehen von den Folgen einer Verletzung des Hippocampus aus, um auf dessen Funktion bei der Übertragung von Informationen ins Langzeitgedächtnis zu schließen. Verletzungen des Hippocampus können einerseits zu anterograder Amnesie führen, d.h. einem Verlust des Langzeitgedächtnisses seit dem Zeitpunkt der Verletzung. Möglich ist aber auch eine retrograde Amnesie, womit ein Verlust des Langzeitgedächtnisses für Erinnerungen vor der Hippocampusläsion bezeichnet wird. Allerdings wird das Kurzzeitgedächtnis nicht beeinträchtigt und auch implizite Gedächtnisleistungen<sup>12</sup>, für die kein bewusster Zugang zum Langzeitgedächtnis benötigt wird, bleiben erhalten. Ausgehend von den Defiziten, die mit einer vollständigen oder teilweisen Zerstörung des Hippocampus verbunden sind, und einer Vielzahl experimenteller Befunde bei Mensch und Tier wird eine Theorie der Funktion des Hippocampus bei der Integration von neuen Informationen in das im Neocortex lokalisierte Langzeitgedächtnis entwickelt. Die Rolle des Hippocampus wird dort als die eines “Lehrers” des Neocortex beschrieben. Seine Aufgabe besteht darin, dem Neocortex neue Informationen wiederholt zu präsentieren. Auf diese Weise gelingt es dem Neocortex innerhalb eines langwierigen Prozesses das neue Wissen endgültig zu speichern (*Konsolidierung*). Durch die Zerstörung des Hippocampus wären damit sowohl anterograde als auch retrograde Amnesie erklärbar: Einerseits fällt mit dem Verlust des “Lehrers” die Fähigkeit weg, neue Informationen ins Langzeitgedächtnis zu integrieren — dies erklärt die anterograde Amnesie. Weil die *Konsolidierung* beim Menschen Jahrzehnte in Anspruch nehmen kann, ergibt sich außerdem die Möglichkeit, die retrograde Amnesie zu erklären: Sie kommt durch den Verlust von Erinnerungen zustande, die noch nicht vom Hippocampus ins Langzeitgedächtnis übertragen worden sind.

### **Der Hippocampus als Pufferspeicher zwischen EG und OES**

In der PSI-Theorie nimmt der Hippocampus die Rolle eines Mittlers zwischen dem Objekterkennungssystem, das neue Informationen wahrnimmt, und dem Extensionsgedächtnis, in das diese neuen Eindrücke als Erfahrungswissen integriert werden sollen, ein. Diese Informationsübertragung ist in Abb. 1.2 dargestellt. Die Aufgabe des Hippocampus besteht nicht nur darin, Wissen durch vielfache Wiederholung zu festigen, sondern auch in der Bereitstellung des Wissens in einem geeigneten Format: Tatsächlich besteht eine weitere Funktion des Hippocampus in der schnellen Bildung weitläufiger Assoziationen. Die im Hippocampus gebildeten Assoziationen sind dem Bewusstsein zugänglich, können verbal beschrieben werden und eignen sich

---

<sup>12</sup>siehe Fußnote 9 auf S.26

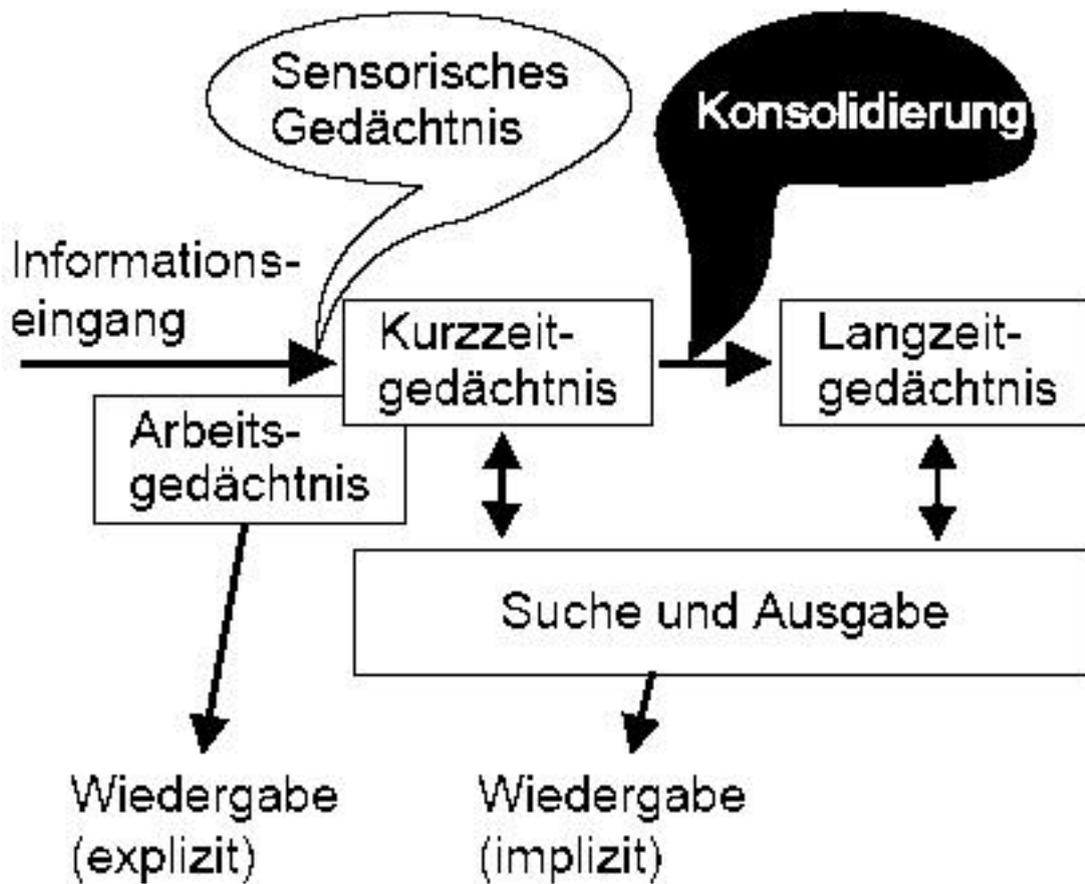


Abbildung 1.1: Die flüchtigen Eindrücke des Kurzzeitgedächtnisses gelangen durch den langwierigen Prozess der Konsolidierung ins Langzeitgedächtnis. Diese Fähigkeit wird in [10, McClelland et al.] dem Hippocampus zugeschrieben.

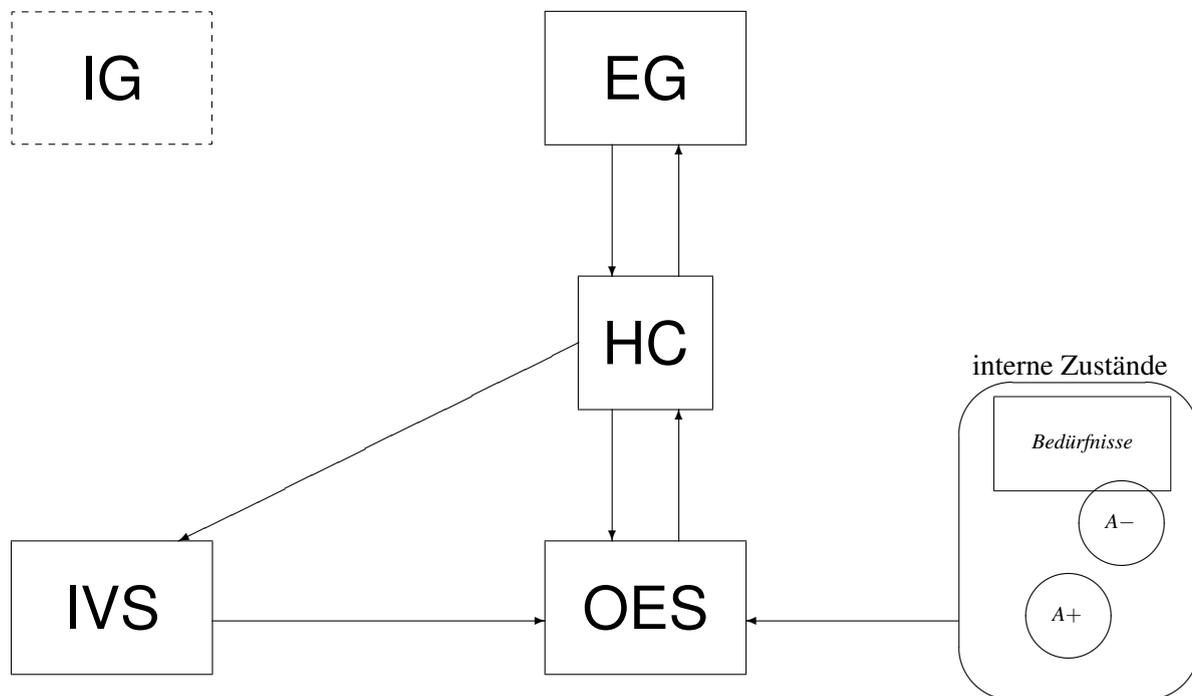


Abbildung 1.2: Hier ist die Informationsübertragung zwischen Extensionsgedächtnis (EG) und den anderen bisher modellierten kognitiven Makrosystemen dargestellt. Ein simulierter Hippocampus (HC), der die in [10, McClelland et al.] beschriebene Rolle spielen soll, überträgt Eindrücke des Objekterkennungssystems (OES) ins Extensionsgedächtnis (EG): Durch das OES wahrgenommene Statusinformationen (Affekt, zuletzt ausgeführte Verhaltensweisen, wahrgenommene Objekte, Befriedigung einzelner Bedürfnisse) werden in binäre Muster konvertiert und durch den HC selektiert. Die ausgewählten Muster werden im EG gespeichert. Bei einer Abfrage des EG (z.B. durch das IVS, um das Verhalten an implizitem Erfahrungswissen auszurichten) übersetzt der HC die binären Pattern wieder in Statusinformationen zurück.

---

zur Steuerung von Verhalten.<sup>13</sup> Kuhl bringt sie mit Tolmans Begriff der *kognitiven Landkarte* in Verbindung [17, Tolman]. Damit werden auch Leistungen des Hippocampus bei der räumlichen Orientierung berücksichtigt. Als Beispiel für die Bildung einer kognitiven Landkarte nennt Kuhl die Fähigkeit, sich auch einen eher unübersichtlichen Gedächtnisinhalt wie die vielen Einzeldrucke einer Party später gegenwärtig zu machen. So ist es möglich, sich nachträglich an einzelne Szenen der Party aus ganz verschiedenen Perspektiven mit Betonung jeweils unterschiedlicher Details zu erinnern. [9, Kuhl, S.493 f.]

### **Eigenschaften der Repräsentation von Informationen im Hippocampus**

Die hohe Spezialisierung der Nervenzellen des Hippocampus wird als Hinweis für die Repräsentation von klar unterscheidbaren Inhalten mit wenig Überschneidungen und — im Vergleich zum Neocortex — unter Beteiligung sehr viel weniger Nervenzellen gesehen. Das Feuereiner Nervenzelle des Hippocampus ist meist an mehrere auslösende Bedingungen gebunden (“conjunctive coding”), was die oben erwähnte Bildung weitläufiger Assoziationen untersttzt. Durch die hohe Spezialisierung der Nervenzellen und die Tatsache, dass nur vergleichsweise wenige Nervenzellen für die Repräsentation verschiedener Inhalte verwendet werden, kommt die geringe Überschneidung der einzelnen Inhalte zustande. Diese führt außerdem dazu, dass Unterschiede betont werden — meistens feuern die Nervenzellen in klar getrennten Gruppen, selbst wenn zwei Inhalte sich tatsächlich in vielen Beziehungen kaum unterscheiden mögen.<sup>14</sup>

### **Die Notwendigkeit schrittweiser Informationsübertragung ins EG**

Die Autoren legen im folgenden besonderen Wert darauf, zu begründen, warum es notwendig ist, dass die Informationsübertragung in den Neocortex *schrittweise* vor sich geht, statt neue Informationen wie in den Speicher eines Computers in einem Schritt “einzuspeichern”. Die Leistung des Langzeitgedächtnisses besteht nicht darin, eine große Anzahl isolierter Fakten zu lernen, sondern diese miteinander in Beziehung zu setzen und so ihre gemeinsame Struktur erkennen zu können. Anhand von neuronalen Netzwerkmodellen, sogenannten Feed-Forward-Netzen (siehe Abschnitt 2.2) illustrieren sie, dass das Erlernen gemeinsamer Struktur nur durch einen schrittweisen und damit langsamen Prozess möglich ist: Wird neues Wissen zu schnell in ein neuronales Netzwerk übertragen, kann es zum Überschreiben, zum “Vergessen”, von bereits gelernten kommen. Erfolgt das Erlernen des neuen Wissens dagegen schrittweise, werden neues und altes Wissen in Beziehung gebracht und können gemeinsam im selben Netzwerk repräsentiert werden (sogenannte *Akkomodation* des Netzwerks mit den neuen Inhalten, siehe [9, Kuhl, S.499]). Analog können die Eindrücke des Objekterkennungssystems also nur durch eine schrittweise Übertragung in das Extensionsgedächtnis integriert werden.

Nachdem die Aufgaben der Einzelsysteme des Fühlens — des Objekterkennungssystems, des Extensionsgedächtnisses und des Hippocampus — im einzelnen geklärt sind, wird nun nach einem geeigneten Modell gesucht, um das Fühlen in ein Computermodell übertragen zu können.

---

<sup>13</sup>“In humans, the hippocampal system appears to be essential for the rapid formation of comprehensive associations among the various elements of specific events and experiences in a form sufficient to sustain an explicit (Schacter, 1987) retrieval of the contents of the experience, so that they can be attested (explicitly recognized as memories), verbally described, or flexibly used to govern subsequent behaviour.” [10, McClelland et al., S.420]

<sup>14</sup>The use of sparse, conjunctive coding in the hippocampus means that its representations of situations which differ only slightly may have relatively little overlap. [10, McClelland et al., S.426]

---

## 2. Neuronale Netzwerkmodelle

Künstliche neuronale Netze sind zu einem Sammelbegriff für sehr unterschiedliche Ansätze geworden. Ihnen ist gemeinsam, dass mehr oder weniger stark vereinfachte “künstliche Nervenzellen” miteinander verbunden werden. Je nach Zielsetzung werden dabei sehr unterschiedliche Typen von Nervenzellen und Verschaltungen (*Architekturen*) gewählt. Eine umfangreiche mathematische Theorie beschreibt Vorzüge und Einschränkungen der einzelnen Architekturen sowie ihre Eignung für unterschiedliche Anwendungen. Dieser Abschnitt dient dazu, die Funktionsweise neuronaler Netzwerkmodelle darzustellen und einige notwendige mathematische Begriffe zu erläutern. Die Informationen zu einzelnen Netzwerkarchitekturen und deren Training wurden größtenteils [3, Hammer] und [21, Zell] entnommen. In 2.3 wird erläutert, wie bestimmte neuronale Netzwerke zur Speicherung von Assoziationen eingesetzt werden können. Mithilfe eines solchen Hopfield-Netzwerks wird später das Extensionsgedächtnis modelliert werden. In Abschnitt 2.4 wird ein Ähnlichkeitsmaß für Assoziationen entwickelt. Dieses Ähnlichkeitsmaß wird benötigt, um die Assoziationen auszuwählen, die in das Extensionsgedächtnis übertragen werden sollen.

### 2.1 Allgemeine Eigenschaften

Künstliche Nervenzellen sind stark vereinfachte Modelle der Nervenzellen des menschlichen Gehirns. Sie haben eine bestimmte Anzahl von Eingängen und einen Ausgang. Durch die Ein-

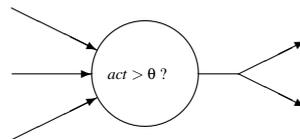


Abbildung 2.3: Schematische Darstellung einer künstlichen Nervenzelle, die Eingaben (im Bild rechts) vorwärts weiterleitet.

gänge erhält eine Nervenzelle Aktivierungen bestimmter Stärke. Die an den einzelnen Eingängen ankommenden Aktivierungen werden durch die Nervenzelle gewichtet und dann aufsummiert. Liegt die Gesamtsumme über einem bestimmten Schwellenwert, wird der Ausgang aktiviert — dies entspricht dem “Feuern” einer natürlichen Nervenzelle — ansonsten bleibt er deaktiviert. Letztendlich berechnet eine Nervenzelle also die Funktion

$$act(v) = \sum_{j=1}^n w_j v_j - \theta. \quad (2.1)$$

Hierbei ist  $v$  der Vektor der an den Eingängen ankommenden Aktivierungen,  $w_j$  sind die Gewichte und  $\theta$  ist der Schwellenwert. Meistens entspricht die Aktivierung der künstlichen Nervenzelle noch nicht ihrer Ausgabe. Diese wird durch Anwendung einer Aktivierungsfunktion  $f$  auf die Aktivierung der Nervenzelle berechnet:

$$o(v) = f(act(v)) = f\left(\sum_{j=1}^n w_j v_j - \theta\right). \quad (2.2)$$

Beispiele für Aktivierungsfunktionen folgen später. Eine einzelne Nervenzelle für sich ist — wie auch ihr natürliches Vorbild im menschlichen Gehirn — noch nicht besonders leistungsfähig. Die Leistungsfähigkeit eines neuronalen Netzes kommt erst durch die Verschaltung vieler

---

Nervenzellen zustande. Immerhin kann man einige boolesche Operatoren (*and* und *or*) bereits mithilfe einer einzigen Nervenzelle berechnen. Doch schon der Operator *xor* ist nur bei Hinzunahme mindestens einer weiteren Nervenzelle darstellbar. Dies kann sehr schön an einem benutzerfreundlichen, im Internet verfügbaren Applet ausprobiert werden (siehe [1]).

Entscheidender Vorteil neuronaler Netze ist jedoch nicht die Tatsache, dass sie die Möglichkeit bieten, bestimmte mathematische Funktionen durch die Bestimmung sehr vieler Parameter, nämlich der Gewichte und Schwellenwerte, eines neuronalen Netzes darstellen zu können — eher im Gegenteil: Dass man Funktionen wie  $f(x) = x^2$  mithilfe eines neuronalen Netzwerkes darstellen kann, ist wenig hilfreich; gegenüber der simplen Funktionsdefinition hat eine Menge von Gewichten und Schwellenwerten, die normalerweise keine Regelmäßigkeit erkennen lassen, zunächst keinen Vorteil. Die entscheidende Stärke neuronaler Netzwerkmodelle besteht darin, dass Gewichte und Schwellenwerte nach standardisierten Verfahren, durch sogenanntes *Training*, berechnet werden können — einzig und allein dadurch, dass dem neuronalen Netz Beispiele für das gewünschte Verhalten gezeigt werden. Durch oft wiederholte Präsentation von Zahlenpaaren wie (1;1), (2;4), (3;9) und (4;16) kann man einem geeigneten neuronalen Netz nicht nur “beibringen”, diese zu reproduzieren, also bei einer Eingabe von 1 an das Netz den Wert 1 zurückzuliefern, bei 2 den Wert von 4 usw. Stattdessen ist es möglich — in gewissen Grenzen, die mathematisch jedoch genau angegeben werden können — einem neuronalen Netzwerk durch solche Beispiele die komplette mathematische Funktionen beizubringen. In unserem Beispiel würde das fertig trainierte Netzwerk oftmals in der Lage sein, auch zu bisher unbekanntem Werten die richtige “Antwort” zu liefern, also gewissermaßen die Funktion  $f(x) = x^2$  zu “erlernen”. (Zur Verdeutlichung wieder ein Applet, siehe [20])

## 2.2 Feedforward-Netze als Funktionsapproximatoren

Der einfachste Art der Verschaltung von Nervenzellen ist ein vorwärtsgerichtetes Netz. Die Nervenzellen sind hier so miteinander verbunden, dass die Aktivierung am Ausgang einer Nervenzelle nur so weitergeleitet wird, dass sie niemals Einfluss auf einen oder mehrere Eingänge derselben Nervenzelle hat. Mathematisch bilden die Neuronen in dieser Art der Verschaltung einen azyklischen, gerichteten Graphen (siehe 2.4). Man kann sich also vorstellen, dass die Aktivierungen der Eingänge der Eingabeneuronen des Netzes über die von außen nicht sichtbaren Zellen in den sogenannten *hidden layers* bis zu den Ausgabeneuronen nur in eine Richtung, also gewissermaßen “vorwärts”, weitergeleitet werden. Die Aktivierung einzelner Nervenzellen kann nun in jeder einzelnen Schicht durch

$$act_j(v) = f \left( \sum_{i=1}^n w_{ij} v_i - \theta_j \right). \quad (2.3)$$

berechnet werden.  $w_{ij}$  ist dabei das Gewicht zwischen dem Neuron mit dem Index  $i$  und dem Neuron mit Index  $j$  in der nächsten Schicht. Dies ist eine Verallgemeinerung der Formel (2.1), da nun nicht nur ein Neuron pro Schicht existiert, sondern mehrere. Für  $v_i$  ist die Aktivierung des  $i$ -ten Neurons der Eingabeschicht einzusetzen. Hierbei ist  $f$  wieder eine Aktivierungsfunktion, für die im Fall eines Feedforward-Netzes die sogenannte Sigmoidfunktion

$$sgd(x) = \frac{1}{1 + \exp(-x)} \quad (2.4)$$

gewählt wird.

Vektoriell ausgedrückt mit einer Gewichtsmatrix  $W = (w_{ij})$  und einem Vektor  $v$ , der die Akti-

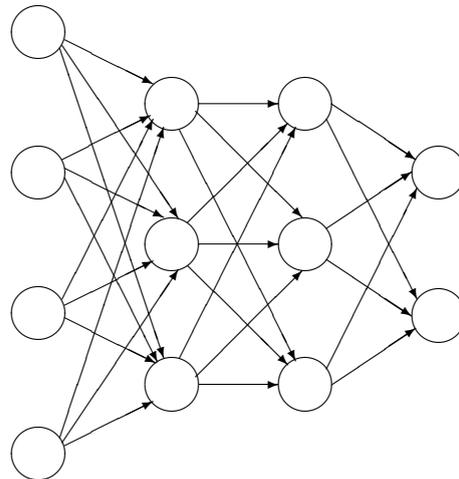


Abbildung 2.4: Feed-Forward-Netze verschalten künstliche Nervenzellen in Form eines vorwärtsgerichteten, azyklischen Graphen.

vierungen der Neuronen der oberen Schicht enthält, wird (2.3) zu

$$O(v) = f(Wv - \Theta). \quad (2.5)$$

Das Standard-Trainingsverfahren *Backpropagation* [11] ermöglicht in vielen Fällen die Berechnung der Gewichte  $w_{ij}$  und der Schwellenwerte  $\theta_j$  in den einzelnen Schichten. Hierauf soll in dieser Arbeit nicht weiter eingegangen werden. Es sei darauf hingewiesen, dass Feed-Forward-Netze theoretisch in der Lage sind, beliebige stetige Funktionen zu lernen. Im PSI-Modell wären sie deshalb einsetzbar, um die Affektdynamik einer Person zu erlernen (siehe Abschnitt 4.). Dies ermöglicht der modellierten Person, aus gegebenen Werten für positiven und negativen Affekt sowie den Aktivierungen der Makrosysteme diese Größen für den nächsten oder die nächsten folgenden Zeitschritte zu prognostizieren. Aufgrund der prognostizierten Aktivierungen könnte das Selbst dann z.B. durch *Selbstmotivierung* (siehe 1.2) auf die Affektlage Einfluss nehmen.

## 2.3 Hopfield-Netze als Assoziativspeicher

Die Assoziationsfähigkeit des Selbst wird durch ein *Hopfield-Netz* modelliert. Diese Architektur wird hier ausführlich erläutert. Außerdem wird in diesem Abschnitt erklärt, welcherart die *Assoziationen* sind, die in diesem Modell dargestellt werden können.

### Hopfield-Netze sind vollständig rekurrente Netze

Ein Hopfield-Netz ist ein vollständig rekurrentes Netz, d.h. alle Nervenzellen sind miteinander verbunden — das Netzwerk ist formal gesehen das kartesische Produkt der Nervenzellen (siehe Abb. 2.5). Wie die einzelnen Schichten des Feed-Forward-Netzes kann auch ein Hopfieldnetz durch eine Gewichtsmatrix dargestellt werden. Jede Komponente  $w_{ij}$  der Matrix enthält ein Gewicht, das die Verknüpfungsstärke zwischen den Nervenzellen  $i$  und  $j$  angibt. Hopfieldnetze sind ein Spezialfall rekurrenter Netze, da die Verknüpfungsstärke zwischen zwei Nervenzellen  $i$  und  $j$  in beiden Richtungen gleich sein muss, d.h. es muss gelten  $w_{ij} = w_{ji}$ . Die Gewichts-

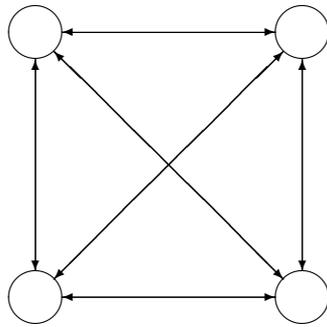


Abbildung 2.5: In einem Hopfield-Netz sind alle Nervenzellen miteinander verbunden. Die Gewichte zwischen den Nervenzellen sind in beiden Richtungen gleich.

matrix ist also symmetrisch. In der Diagonalen stehen jeweils die Gewichte, mit denen eine Nervenzelle sich selbst aktiviert.

### Darstellung von Assoziationen im Hopfield-Netz

Ein Hopfield-Netz kann nur binäre Aktivierungen verarbeiten, d.h. die beiden Werte -1 und 1. Die Aktivierung 1 steht dabei für eine aktivierte, die Aktivierung -1 für eine deaktivierte Nervenzelle. Die Frage ist nun, wie mit binären Aktivierungen einzelner Nervenzellen Assoziationen dargestellt werden können. Der Einfachheit halber wird zunächst davon ausgegangen, dass die Situationsmerkmale, zwischen denen Assoziationen gebildet werden sollen, nur die Werte “vorhanden” bzw. “nicht vorhanden” einnehmen können. Dann kann das Vorhandensein jedes Situationsmerkmal durch die Aktivierung jeweils einer Nervenzelle dargestellt werden. Eine Assoziation zwischen verschiedenen Situationsmerkmalen liegt dann vor, wenn die für die Merkmale zuständigen Merkmale gleichzeitig aktiviert sind. Im folgenden Beispiel sollen Assoziationen zwischen den Affekten ( $A+$  und  $A-$ ), Annäherungs- bzw. Meidungsverhalten ( $Ann$  und  $Meid$ ) und zwei verschiedenen Sorten von Objekten ( $000$  und  $111$ ) dargestellt werden.



In der Zeichnung stehen ausgefüllte Kreise für aktivierte und nicht ausgefüllte Kreise für deaktivierte Nervenzellen — die Verbindungen zwischen den Zellen sind aus Gründen der Übersichtlichkeit nicht eingezeichnet, da es hier nur auf ihren Zustand ankommt. Das aus sechs Nervenzellen bestehende Netzwerk zeigt also eine Assoziation zwischen positivem Affekt, Annäherungsverhalten und einem Objekt der Kategorie  $000$  an — die deaktivierten Nervenzellen sind so zu verstehen, dass die betreffenden Situationsmerkmalen mit den assoziierten Merkmalen nicht in Verbindung stehen. Mit diesem Modell können demnach Assoziationen der Form



---

ausgedrückt werden — hier sind einerseits positiver Affekt, Annäherungsverhalten und ein Objekt der Kategorie 000 bzw. negativer Affekt und ein Objekt der Kategorie 111 miteinander assoziiert. Allerdings kann dieses Modell bei Bedarf leicht verfeinert werden. Dies kann zum Beispiel sinnvoll sein, wenn Assoziationen mit Bedürfnissen gebildet werden sollen. Hier ist es wichtig, zumindest drei unterschiedliche Zustände zu unterscheiden: Ein Bedürfnis kann entweder befriedigt oder unbefriedigt sein oder im Moment eine Mittelstellung einnehmen, in der es weder an der Generierung positiver noch negativer Affekte beteiligt ist. Dies kann mit zwei Nervenzellen dargestellt werden: Eine steht für die Befriedigung, eine andere für die Frustration des Bedürfnisses. Ein mittleres Aktivierungsniveau kann z.B. durch die gleichzeitige Deaktivierung beider Nervenzellen modelliert werden (weder befriedigt, noch frustriert; die gleichzeitige Aktivierung beider Nervenzellen macht in diesem Beispiel natürlich keinen großen Sinn). Auf diese Weise erreicht man für die meisten Anwendungen durch Hinzufügen einer ausreichenden Anzahl von Nervenzellen eine genügend genaue Darstellung der Ausprägung eines Situationsmerkmals.

### Schaltdynamik eines Hopfield-Netzes

Beim Hopfieldnetz erscheint es im Vergleich mit dem Feed-Forward-Netz zunächst als ungewöhnlich, dass Ausgabe- und Eingabeneuronen einander entsprechen. Tatsächlich wird die Ausgabe eines Hopfieldnetzes etwas anders berechnet als bei Feed-Forward-Netzen — sie ähnelt eher der Abfrage eines Speichers als der Berechnung einer Funktion: Ein Eingabevektor  $v$  wird dem Netz präsentiert und durch Multiplikation mit der Matrix  $W$  der Verknüpfungsgewichte eine Ausgabe  $O$  berechnet:

$$O(v) = f(Wv) \quad (2.6)$$

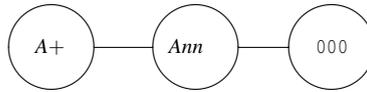
Ein Schwellwertvektor  $\Theta$  wird für unsere Anwendung nicht benötigt. Für die Aktivierungsfunktion wird die in (2.7) komponentenweise ausgeführte Aktivierungsfunktion verwendet. Für jede Komponente  $i$  gilt:

$$f_i(v_i) = \begin{cases} 1, & \text{für } v_i \geq 0 \\ -1, & \text{für } v_i < 0 \end{cases} \quad (2.7)$$

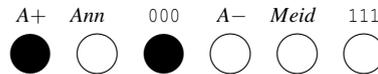
Die Gleichung (2.6) ähnelt der Gleichung (2.5) für Feed-Forward-Netze. (2.6) ist jedoch eine Rekursionsformel: Das Ergebnis von (2.6), also der neue Zustand der Nervenzellen, wird mit dem Eingabevektor verglichen. Einerseits kann er unverändert zurückgegeben werden — dies ist der Fall, wenn der Eingabevektor zuvor im Netz *gespeichert* worden ist. Stimmen Ein- und Ausgabe nicht überein, wird der Ausgabevektor nochmals gemäß (2.6) in das Netz eingegeben und erneut mit der neuen Ausgabe verglichen. Dies wird so lange wiederholt, bis der eingegebene Vektor dem ausgegebenen Vektor entspricht, d.h. bis also gilt:  $O(v) = v$ . Dieses Verhalten wirkt auf den ersten Blick weniger interessant als die Fähigkeit von Feed-Forward-Netzen, mit denen immerhin mathematische Funktionen approximiert werden können. Die Stärke eines Hopfieldnetzes liegt jedoch darin, Eingabevektoren zu möglichst “ähnlichen” Ausgabevektoren, die im Netz gespeichert wurden, zu ergänzen. Dies kann als einfaches Modell für ein assoziatives Gedächtnis interpretiert werden, wie das folgende Beispiel zeigen soll.

### “Assoziieren” mit einem Hopfield-Netz

Ohne zunächst zu wissen, wie dies geschieht, nehmen wir an, dass in dem obigen Beispiel die Assoziation



im Hopfield-Netzwerk “gespeichert” ist, d.h. dass sie nach (2.6) durch Präsentation genau dieser Assoziation ( $A+$  aktiv,  $Ann$  aktiv,  $Lust$  aktiv,  $A-$  inaktiv,  $Meid$  inaktiv,  $Furcht$  inaktiv, d.h. dem Vektor  $(1,1,1,-1,-1,-1)$ ) abgerufen werden kann. Zumindest wenn nur diese Assoziation gespeichert ist, liefert auch die Eingabe



die Netzantwort



Das Hopfield-Netz “assoziert” also zu positivem Affekt und einem Objekt der Kategorie 000 Annäherungsverhalten. Ebenso kann zu Annäherungsverhalten und dem Objekt 000 positiver Affekt “assoziert” werden. Diese Assoziationsfähigkeit, die in der Ergänzung noch “fehlender” Information zu den präsentierten Situationsmerkmalen besteht, wird später zur Steuerung des Verhaltens der simulierten Person eingesetzt (siehe 3.5).

## Training

Zunächst stellt sich aber noch die Frage, wie Assoziationen der Art, wie sie oben vorgestellt wurden, in einem Hopfield-Netz gespeichert werden können, wie Hopfield-Netze also trainiert werden. Das für unser Modell verwendete Trainingsverfahren ist deutlich einfacher als der *Backpropagation*-Algorithmus. Zwar können Hopfield-Netze auch mit dem *Perzeptron*-Algorithmus trainiert werden. Dieser Algorithmus ähnelt dem *Backpropagation*-Algorithmus, da auch beim *Perzeptron*-Algorithmus aus den Fehlern die Korrekturwerte für die Gewichte des Netzes bestimmt werden. Dies ist jedoch hier nicht notwendig, die Gewichte können mithilfe der *Hebb-schen Regel* (2.8) direkt berechnet werden:

$$w_{ij} = \frac{1}{n} \sum_{p=1}^n X_i^p X_j^p \quad (2.8)$$

Oft werden die Gewichte  $w_{ii}$  auf 0 gesetzt. Damit wird verhindert, dass Nervenzellen sich selbst aktivieren können. In diesem Modell werden jedoch auch Selbstaktivierungen der Nervenzellen zugelassen.

## Ein Beispiel für ein Hopfieldnetzwerk

Betrachten wir als Beispiel die Assoziation



Dies entspricht dem Vektor

$$v = \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}$$

Dieser Vektor soll nun als einziges Muster in einem Hopfieldnetz gespeichert werden. Nach (2.8) erhält man die Matrix

$$W = \begin{pmatrix} 0.25 & 0.25 & -0.25 & -0.25 \\ 0.25 & 0.25 & -0.25 & -0.25 \\ -0.25 & -0.25 & 0.25 & 0.25 \\ -0.25 & -0.25 & 0.25 & 0.25 \end{pmatrix}. \quad (2.9)$$

Tatsächlich erkennt man, dass der Vektor  $v$  im Hopfield-Netz “gespeichert” ist, denn man erhält für die Ausgabe  $O$  des Netzes:

$$O(v) = Wv = \begin{pmatrix} 0.25 & 0.25 & -0.25 & -0.25 \\ 0.25 & 0.25 & -0.25 & -0.25 \\ -0.25 & -0.25 & 0.25 & 0.25 \\ -0.25 & -0.25 & 0.25 & 0.25 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix} \quad (2.10)$$

und auch das komponentenweise Anwenden von (2.7) ändert nichts an diesem Ergebnis. Der Vektor  $v$  ist also im Hopfieldnetz gespeichert, denn wird er dem Netz als Eingabe präsentiert, wird er unverändert zurückgeliefert.

Weit interessanter ist jedoch, dass das Netz ebenfalls

$$\begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}$$

zurückgibt, wenn

$$v' = \begin{pmatrix} 1 \\ -1 \\ -1 \\ -1 \end{pmatrix}$$

eingegeben wird:

$$o(v') = Wv' = \begin{pmatrix} 0.25 & 0.25 & -0.25 & -0.25 \\ 0.25 & 0.25 & -0.25 & -0.25 \\ -0.25 & -0.25 & 0.25 & 0.25 \\ -0.25 & -0.25 & 0.25 & 0.25 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0.5 \\ 0.5 \\ -0.5 \\ -0.5 \end{pmatrix} \quad (2.11)$$

und nach komponentenweisem Anwenden von (2.7) erhält man den Vektor  $v$ . Man könnte also sagen, dass das Netz eine Assoziation zwischen positivem Affekt und einem Objekt der Kategorie  $000$  hergestellt hat, denn bei einer Eingabe, die nur positiven Affekt enthält, wird in der

Ausgabe zusätzlich die Objektkategorie 000 aktiviert. Umgekehrt führt die Eingabe, in der ausschließlich die Objektkategorie 000 aktiviert ist, zu einer Ausgabe, in der auch positiver Affekt auftaucht.

## 2.4 Ähnlichkeit von Assoziationen

Für beliebige, mit zufälligen Werten initialisierte Muster haben Hopfield-Netze eine sehr geringe Speicherkapazität. Man kann zeigen, dass sie nur  $0.14N$  Muster speichern können, wobei  $N$  die Anzahl der Nervenzellen des Hopfield-Netzes ist. Es ist also erforderlich, eine Methode zu entwickeln, nach der die zu speichernden Assoziationen gemäß den in Abschnitt 1.5 erläuterten Eigenschaften des Hippocampus ausgewählt werden.

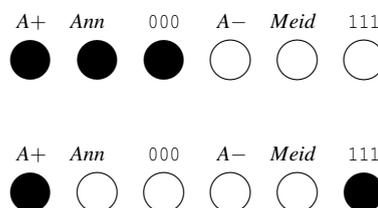
### Auswahl von Assoziationen zur Übertragung ins EG

Nach Abschnitt 1.5 sollten solche Muster zur Übertragung ins Extensionsgedächtnis ausgewählt werden, die sich möglichst stark unterscheiden. Um dies mathematisch abbilden zu können, betrachten wir die Multiplikationstabelle 3.1. Fasst man wie im Abschnitt über die Assoziations-

**Tabelle 3.1: Multiplikationstabelle für die Zahlen -1 und 1.**

·	1	-1
1	1	-1
-1	-1	1

fähigkeit von Hopfield-Netzen (siehe Abschnitt 2.3) den Wert 1 als das Vorhandensein und -1 als das Nicht-Vorhandensein eines Merkmals auf, so drückt die Multiplikation die Übereinstimmung einzelner Merkmale einer Assoziation aus. Wir erläutern dies an einem Beispiel (siehe Abb. 2.6): Die Abbildung zeigt zwei Assoziationen, die nur darin übereinstimmen, dass sie po-



*Abbildung 2.6: Bei binären Assoziationen “zählt” das Skalarprodukt die Übereinstimmungen. Hier stimmen die zwei Assoziationen nur darin überein, dass sie beide jeweils positiven Affekt enthalten. Da auch Merkmale, die in beiden Assoziationen nicht vorkommen, berücksichtigt werden, ergibt sich für die hier gezeigten Assoziationen nach (2.15) der Wert 0.*

sitiven Affekt (A+) enthalten. Stellt man die beiden Assoziationen als Vektoren dar, erhält man für die Assoziation von positivem Affekt (A+), Annäherungsverhalten (Ann) und der Objektkategorie 000 den Vektor (1,1,1,-1,-1,-1) und für die Assoziation von negativem Affekt (A–)

und dem Objekt der Kategorie 111 den Vektor  $(1,-1,-1,-1,-1,1)$ . Durch Multiplikation der einzelnen Komponenten kann jeweils festgestellt werden, ob zwei Komponenten übereinstimmen — dann erhält man den Wert 1 — oder nicht übereinstimmen — in diesem Fall erreicht man -1. In unserem Beispiel ergeben sich drei Übereinstimmungen und drei Nicht-Übereinstimmungen: Die beiden Assoziationen stimmen darin überein, dass sie jeweils  $A+$  enthalten, allerdings auch darin, dass sie *nicht*  $A-$  und *kein* Meidungsverhalten enthalten. Ob die Anzahl der Übereinstimmungen oder die Anzahl der Nicht-Übereinstimmungen überwiegt, kann man einfach “zählen”, indem man die Produkte der einzelnen Komponenten zweier Vektoren  $v$  und  $w$  addiert. Dies entspricht einem wohlbekanntem Begriff der Vektorrechnung, dem *Skalarprodukt*:

$$\langle v, w \rangle := \sum_{i=1}^n v_i w_i. \quad (2.12)$$

Für unser Beispiel erhält man für  $v = (1, 1, 1, -1, -1, -1)$  und  $w = (1, -1, -1, -1, -1, 1)$ :

$$\langle v, w \rangle = 1 \cdot 1 + 1 \cdot (-1) + 1 \cdot (-1) + (-1) \cdot (-1) + (-1) \cdot (-1) + (-1) \cdot 1 = 0.$$

Der Wert 0 drückt aus, dass die Vektoren in genauso vielen Komponenten übereinstimmen wie nicht übereinstimmen; wenn die Mehrzahl der Komponenten übereinstimmt, ergibt sich ein positiver Wert, stimmt die Mehrzahl der Komponenten beider Vektoren nicht überein, ergibt sich ein negativer Wert. Mit

$$U(v, w) := \text{Anzahl der Übereinstimmungen von } v \text{ und } w \quad (2.13)$$

$$N(v, w) := \text{Anzahl der Unterschiede von } v \text{ und } w \quad (2.14)$$

kann man also auch schreiben:

$$\text{Overlap}(v, w) := \langle v, w \rangle = U(v, w) - N(v, w). \quad (2.15)$$

Das Skalarprodukt ist also ein Maß für die Überschneidungen, den “Overlap”, von binären Vektoren in den Einträgen -1 und 1. Interessieren wir uns nach Abschnitt 1.5 besonders für Paare von Assoziationen, die möglichst wenig übereinstimmen, sind negative Werte des Skalarprodukts von besonderer Wichtigkeit: Je kleiner der Wert des “Overlaps”, um so weniger stimmen die Assoziationsmuster überein.

### Der Winkel als geometrische Veranschaulichung der Ähnlichkeit

Die Ähnlichkeit von Assoziationen kann mit der geometrischen Interpretation des Skalarprodukts auch sehr gut bildlich veranschaulicht werden: “Ähnlichkeit” im Sinne von (2.15) entspricht der Größe des Winkels zwischen den Vektoren, die zu den Assoziationen gehören (siehe Abb. 2.7). Zur Berechnung des Winkels wird die Länge der beiden Vektoren benötigt (siehe Abb. 2.8). Die Länge  $\|v\|$  eines Vektors  $v$  kann mithilfe des Skalarprodukts berechnet werden:

$$\|v\| = \sqrt{\langle v, v \rangle} = \sqrt{\sum_{i=1}^n v_i v_i} = \sqrt{\sum_{i=1}^n v_i^2}. \quad (2.16)$$

Der Winkel  $\alpha$  zwischen den beiden Vektoren  $v$  und  $w$  ergibt sich aus dem Skalarprodukt und den Längen  $\|v\|$  und  $\|w\|$ :

$$\cos \alpha = \frac{\langle v, w \rangle}{\|v\| \|w\|} \quad (2.17)$$

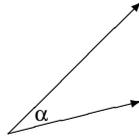


Abbildung 2.7: Für Assoziationsmuster ist der Winkel  $\alpha$  zwischen den beiden zugehörigen binären Vektoren ein Maß dafür, wie “ähnlich” die beiden Assoziationen sind, d.h. wie stark sie sich überlappen: Ähneln die Assoziationen sich stark, erhält man einen kleinen Winkel in der Nähe von 0 Grad. Sind die Vektoren dagegen sehr unterschiedlich, bilden sie einen großen Winkel in der Nähe von 180 Grad.



Abbildung 2.8: Vektoren können geometrisch als Pfeile veranschaulicht werden. Ein Vektor weist dann in eine bestimmte Richtung und hat eine Länge  $||v||$

Die rechte Seite dieser Gleichung ist nur eine Normierung von (2.15), der Bruch nimmt Werte zwischen 1 und -1 an. Damit sind die Voraussetzungen erfüllt, Vektoren beliebiger Länge auf ihre Ähnlichkeit zu vergleichen: Für “ähnliche” Vektoren hat der Bruch auf der rechten Seite von (2.17) einen Wert in der Nähe von 1, was einem kleinen Winkel in der Nähe von 0 Grad entspricht. Stark unterschiedliche Vektoren liefern auf der rechten Seite einen Wert in der Nähe von -1 — dies entspricht einem großen Winkel in der Nähe von 180 Grad. Unser Beispiel aus Abb. 2.6 liegt mit einem Winkel von 90 Grad genau in der Mitte, ist für eine Speicherung im Extensionsgedächtnis also nicht unbedingt zu empfehlen, aber auch nicht völlig unbrauchbar. Da der Winkel in erster Linie zur geometrischen Veranschaulichung diente, reicht es für unsere

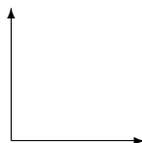


Abbildung 2.9: Die Assoziationen des Beispiels aus Abb. 2.6 stehen als Vektoren in einem Winkel von 90 Grad zueinander, nach (2.18) haben sie ein Ähnlichkeitsmaß von 0. Dies liegt genau in der Mitte; insgesamt reichen die Werte von -1 bis 1, dem entsprechen nach (2.17) Winkel von 0 bis 180 Grad.

Zwecke aus, Assoziationen zukünftig nur mithilfe der rechten Seite von (2.17) zu vergleichen. Hier kann noch ausgenutzt werden, dass binäre Vektoren stets eine Länge  $||v|| = \sqrt{n}$  haben, so dass man vereinfacht erhält:

$$\sigma(v, w) = \frac{1}{n} \langle v, w \rangle \tag{2.18}$$

### Verallgemeinerung für mehr als zwei Muster

Bisher wurde stets nur die Ähnlichkeit zweier Vektoren betrachtet. Da das Extensionsgedächtnis viele Muster speichern soll, müssen die in den vorangegangenen Abschnitten entwickelten Ähnlichkeitsmaße noch für mehrere Vektoren verallgemeinert werden. Dies ist recht einfach:

---

Für eine beliebige Menge von Vektoren wird der Durchschnitt aus den einzelnen Maßen  $\sigma$  für alle Paare von Vektoren berechnet:

$$\Omega(\{v_1, \dots, v_n\}) = \frac{2}{n(n-1)} \sum_{i=1}^n \sum_{j=i}^n \sigma(v_i, v_j), \quad (2.19)$$

wobei für  $\sigma$  das Maß (2.18) eingesetzt wird. Der Vorfaktor  $\frac{2}{n(n-1)}$  dient zur Normierung, er teilt durch die Anzahl  $\frac{n(n-1)}{2}$  der Summanden in (2.19), um Mengen von Mustern mit unterschiedlich vielen Elementen miteinander vergleichen zu können. Mit (2.19) erhält man somit ein Maß, mit dem beliebige Mengen von Assoziationen auf ihre Eignung für eine Übertragung ins Extensionsgedächtnis geprüft werden können.

### 3. Modellbeschreibung

In diesem Abschnitt sollen die im vorigen Abschnitt entwickelten Einzelbestandteile eines Modells des Fühlens — das Extensionsgedächtnis als Hopfield-Netzwerk und der Hippocampus als “Trainer” dieses Netzwerkes, der aus Wahrnehmungen des Objekterkennungssystems Assoziationen bildet und möglichst unterschiedliche ins Extensionsgedächtnis überträgt — zusammengesetzt und auf ihre Konsistenz mit den Annahmen der PSI-Theorie untersucht werden. Eine experimentelle Überprüfung des Modells findet am Schluss dieses Kapitels statt.

#### 3.1 Das OES als Lieferant von Eindrücken

Das Objekterkennungssystem muss in diesem Modell keine besonders hohen Anforderungen erfüllen. Seine Aufgabe besteht darin, in jedem Zeitschritt *innere Zustände* wie den Status der Bedürfnisse, die Affekte und ein möglicherweise ausgeführtes Verhalten zu protokollieren; außerdem ein *Objekt der Umgebung* auszuwählen, das als Gegenstand der gegenwärtigen Aufmerksamkeit interpretiert werden kann (durch dieses ausgewählte Objekt wird der *antizipierte Affekt* (siehe Abschnitt 4.3) berechnet). Dieses *Zustandsprotokoll* (siehe Abb. 3.10) der inneren Zustände und äußeren Eindrücke wird dann zur Weiterverarbeitung an den Hippocampus übergeben. Mit den Annahmen der PSI-Theorie über das Objekterkennungssystem stimmt hier

$A+ : 1.03, A- : 0.97, \text{ Verhalten: Annäherung (Ann), Objekt: 000}$

*Abbildung 3.10: Das OES fertigt für jeden Zeitschritt ein Zustandsprotokoll an. Es werden die Affekte, ein möglicherweise ausgeführtes Verhalten und die Objekte protokolliert, die Gegenstand der Aufmerksamkeit der Person waren.*

die *Flüchtigkeit* der Zustandsprotokolle überein: Schon im nächsten Zeitschritt wird der Zustand wieder überschrieben und falls er nicht vom Hippocampus in das Extensionsgedächtnis übertragen wird, “vergessen”.

#### 3.2 Der Hippocampus

Die Merkmale des Hippocampus werden in erster Linie durch das Ähnlichkeitsmaß aus Abschnitt 1.5 modelliert. Da der Hippocampus für die Generierung der Assoziationen zuständig

---

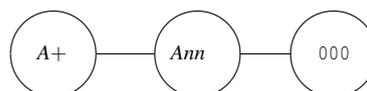
ist, die ins Extensionsgedächtnis übertragen werden, wird noch einmal genauer als im Abschnitt 2.3 auf den Assoziationsbegriff eingegangen, der dem Modell zugrundeliegt. Es wird festgestellt, dass die Art, wie Erfahrungen gesammelt werden, den Vorstellungen eines radikalen Empirismus nach den Vorstellungen des Philosophen David Hume entspricht. Die Konsequenzen werden im Abschnitt 3.5 über die Verhaltenssteuerung durch das Extensionsgedächtnis weiter untersucht.

### Selbstrepräsentationen und integrierte Selbstrepräsentationen

Nach Abschnitt 1.4 enthält das Extensionsgedächtnis in erster Linie autobiographisches Erfahrungswissen, wovon die sogenannten Selbstrepräsentationen eine Sonderstellung einnehmen, da sie einen wesentlichen Teil der Identität einer Person ausmachen. Hier wird davon ausgegangen, dass ein Zustandsprotokoll der Art wie es das Objekterkennungssystem liefert, bereits eine Selbstrepräsentation ist: Die Person bildet ihre *internen Zustände* ab, repräsentiert sich also "selbst". Allerdings sind diese Selbstrepräsentationen der Objekterkennung gemäß der Verarbeitungscharakteristik des Objekterkennungssystems isoliert, ohne Beziehung zu anderen Selbstrepräsentationen. Gerade dies unterscheidet die Selbstrepräsentationen des Objekterkennungssystems von den *integrierten Selbstrepräsentationen* des Extensionsgedächtnisses. Integrierte Selbstrepräsentationen sind in das assoziative Netzwerk des Extensionsgedächtnisses eingebettet. Die große Leistung des Hippocampus besteht in seinem Potential, isolierte Selbstrepräsentationen durch Übertragung in das Extensionsgedächtnis in integrierte Selbstrepräsentationen umzuwandeln.

### Assoziationen

Die Fähigkeit des Hippocampus, Assoziationen zu bilden, besteht im vorliegenden Modell schlicht darin, dass ein Zustandsprotokoll des Objekterkennungssystems automatisch als Assoziation aufgefasst wird. Unter einer *Assoziation* wird im folgenden also nichts als eine Verknüpfung zwischen einzelnen Situationsmerkmalen verstanden, deren Zusammentreffen zu einem bestimmten Zeitschritt festgehalten worden ist. Beispielsweise möge das Objekterkennungssystem auf diese Weise positiven Affekt (*A+*), ein bedürfnisbefriedigendes Objekt (*Lust*) und Annäherungsverhalten (*Ann*) in Zusammenhang bringen, d.h. die Assoziation



kommt nach der Annäherung an ein Objekt der Kategorie 000 zustande, wobei positiver Affekt vorhanden war. Verknüpft eine Person diese drei isolierten Informationen über den inneren Zustand (positiver Affekt und das im letzten Zeitschritt ausgeführte Annäherungsverhalten) und die Wahrnehmung der Außenwelt (Objekt 000) assoziativ miteinander, führt zukünftig das Richten der Aufmerksamkeit auf nur eines der drei Situationsmerkmale zu Aufmerksamkeit für die anderen beiden Merkmale. So würde die Vorbereitung von Annäherungsverhalten an Objekte der Kategorie 000 positiven Affekt auslösen, Annäherung an das Objekt 000 bekommt also eine positive affektive Bewertung. Umgekehrt wird nun auch positiver Affekt mit Annäherungsverhalten und Objekten der Kategorie 000 in Zusammenhang gebracht.

---

## Assoziationsbildung aus der Sicht des Empiristen David Hume

Die modellierte Person verhält sich damit ganz nach den Vorstellungen eines radikalen Empirismus: Der Philosoph David Hume führte das Bilden von Assoziationen auf die drei Prinzipien *Ähnlichkeit*, *Berührung in Raum und Zeit* und *Ursache-Wirkungs-Beziehungen* zurück<sup>15</sup>; die hier gebildeten Assoziationen entsprechen also in der Humeschen Terminologie denen der zweiten Art. Unsere simulierte Person entpuppt sich damit als ein "Humesches Wesen", das sich — zumindest in diesem Teil des Modells — *ausschließlich* auf Erfahrungswissen verlässt, ohne dieses durch vorbereitete Mechanismen einer wie auch immer gearteten "Vernunft" einer Prüfung zu unterziehen. Ein so starkes Vertrauen in "Erfahrung", die hier zudem nur aus einem mehr oder weniger mechanischen Kopieren der wahrgenommenen Reize in die tabula rasa des Gedächtnisses besteht, ist in der philosophischen Diskussion äußerst umstritten. Es scheint tatsächlich kaum vorstellbar, dass in einem bloßen Aufzeichnen des eigenen Zustands und der äußeren Eindrücke "Wissen" entsteht, dass für eine Person als "Erfahrung" von Nutzen sein könnte. Eine gewisse Strukturierung der Zustandsprotokolle des Objekterkennungssystem findet jedoch immerhin durch den Selektionsprozess des Hippocampus statt. Ob dieser komplexe, aber doch relativ mechanistische Prozess ausreichend ist, um die Entstehung des *Fühlens* vollständig ohne Rückgriff auf Prozesse außerhalb der Erfahrung zu erklären, erfordert allerdings eine ausführliche Überprüfung des Modells, die über den Rahmen dieser Arbeit hinausgeht.

### Informationsreduktion

Um diejenigen Assoziationen auszuwählen, die ins Hopfield-Netz des Extensionsgedächtnisses übertragen werden sollen, werden die Assoziationen in binäre Vektoren mit den Einträgen -1 und 1 umgewandelt. In diesem Vektor sind für jedes bekannte Zustandsmerkmal ein oder mehrere Bits reserviert. Bits für im Assoziationsmuster auftauchenden Zustandsmerkmale werden auf den Wert 1 gesetzt, die anderen auf -1. Liegt das Niveau des positiven Affekts über dem Ruhenniveau 1 wird das Bit für positiven Affekt gesetzt, ansonsten erhält es den Wert -1. Die Umwandlung in ein Bitmuster, in dem *alle bekannten* Zustandsmerkmale auftauchen, bewirkt, dass die Assoziation bereits an dieser Stelle mit dem gesamten bereits bekannten Wissen in Zusammenhang gebracht wird, allerdings nur in der Form, dass die nicht zur Assoziation gehörigen Merkmale durch deaktivierte Bits dargestellt werden. Eine Assoziation von positivem Affekt, Annäherungsverhalten und der Objektkategorier 000 erhält also die folgende Form, wie sie bereits in Abschnitt 2.3 vorgestellt wurde:



Indem nur berücksichtigt wird, ob einzelne Merkmale zutreffen oder nicht, wird Information reduziert. In dieser Hinsicht verhält sich das Modell des Hippocampus entsprechend seinem natürlichen Vorbild. Wie stark diese Informationsreduktion ist, hängt davon ab, wie viele Bits zur

---

<sup>15</sup>“Für mich ergeben sich nur drei Prinzipien der Vorstellungsverknüpfung, nämlich *Ähnlichkeit* (*resemblance*), raumzeitliche *Berührung* (*contiguity*) und *Ursache* oder *Wirkung* (*cause or effect*). Daß diese prinzipien der Vorstellungsverknüpfung dienen, wird — wie ich meine — kaum bezweifelt werden. Ein Bild lenkt unsere Gedanken naturgemäß auf das Original; die Erwähnung eines der Wohnräume in einem Gebäude bringt selbstverständlich die Frage oder das Gespräch auf die anderen; und wenn wir an eine Wunde denken, können wir den Gedanken an den ihr folgenden Schmerz kaum vermeiden.” [5, Hume, S.39 f.]

Codierung für ein Merkmal verwendet werden (siehe Abschnitt 2.3: “Assoziieren” mit einem Hopfield-Netz).

### Hervorhebung von Unterschieden

Bei der Auswahl der Assoziationen, die ins Extensionsgedächtnis übertragen werden sollen, werden Unterschiede hervorgehoben, wie bereits in Abschnitt 2.4 erläutert wurde. Dies geschieht, um eine redundante Menge von Assoziationen wie die in Abb. 3.11 gezeigte auszuschließen. Speichert man nicht nur die Assoziation zwischen positivem Affekt, Annäherungs-

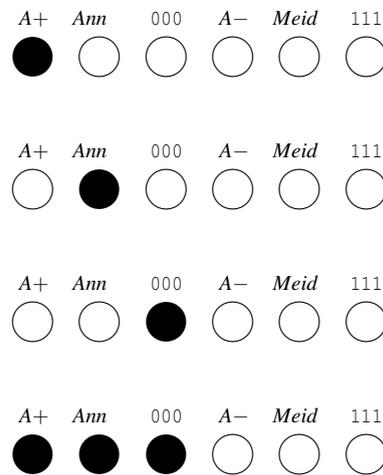


Abbildung 3.11: Ein Beispiel für eine redundante Menge von Assoziationen. Wird diese Menge in ein Hopfield-Netz übertragen, wird die einzige “wirkliche” Assoziation bestehend aus A+, Ann und 000 “verdeckt”: Das Netz kann A+, Ann und 000 nicht mehr zur Assoziation bestehend aus A+, Ann und 000 ergänzen.

verhalten und der Objektkategorie 000, sondern alle drei Bestandteile zusätzlich einzeln, verliert das Hopfield-Netz seine Fähigkeit, wie in Abschnitt 2.3, vorgestellt, zu “assoziieren”: Präsentiert man dem Netz einen Teil der Assoziation, wird dieser nicht mehr ergänzt, da ja auch die einzelnen Bestandteile im Netz gespeichert sind. Solche Fehler können durch das Ähnlichkeitsmaß (2.19) verhindert werden: Finden sich im Pufferspeicher des Hippocampus noch andere

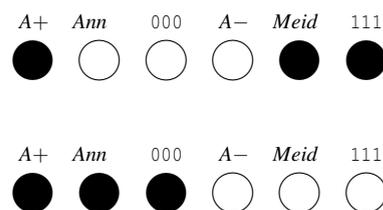
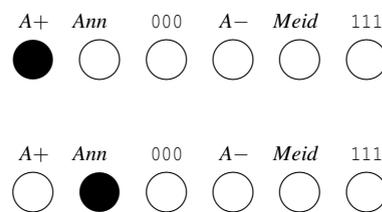


Abbildung 3.12: Ist im Hippokampus zusätzlich zu der Assoziation von positivem Affekt, Annäherungsverhalten und der Objektkategorie 000 auch eine Verknüpfung von positivem Affekt, Meidungsverhalten und der Kategorie 111 vorhanden, werden diese beiden Muster eher zusammen ins Extensionsgedächtnis übertragen als die in Abb. 3.11 gezeigten.

Assoziationen, die weniger Überschneidungen aufweisen, wie z.B. die in Abb. 3.12 gezeigten, so werden diese ins Extensionsgedächtnis übertragen. Damit bevorzugt der Hippocampus zu bekanntem Erfahrungswissen diskrepante Assoziationen, sobald er sie vom Objekterkennungssystem erhält.

### Bildung weitläufiger Assoziationen

Möglicherweise verwunderlich mag am Ähnlichkeitsmaß (2.19) erscheinen, dass es auch das Nicht-Vorhandensein von Merkmalen in zwei Assoziationen als Übereinstimmung wertet. Dies zeigt sich schon an unserem Einführungsbeispiel zum Ähnlichkeitsmaß, das in Abb. 2.6 dargestellt ist: Der Wert von 0 ergab sich daraus, dass die Assoziationen neben dem bei beiden vorhandenen positiven Affekt darin “übereinstimmten”, dass sie beide weder die Merkmale *A-* noch *Meid* enthielten. Dies erscheint auf den ersten Blick merkwürdig, da die Beurteilung nun davon abhängt, welche Zustandsmerkmale dem Hippocampus *außer* den in den Assoziationen vorkommenden noch bekannt sind. Damit werden auch Mengen von Assoziationen, die wenig Merkmale miteinander verknüpfen, schlecht bewertet, auch wenn sie als Assoziationen nicht übereinstimmen (siehe Abb. 3.13). Sollen beispielsweise die isolierten Merkmale *A+* und *Ann*



*Abbildung 3.13: Assoziationen, in denen wenig Merkmale miteinander verknüpft sind, werden durch das Ähnlichkeitsmaß (2.18) schlecht bewertet. Zwar stimmen die Assoziationen an keiner Stelle überein, dafür assoziieren sie keine Zustandsmerkmale miteinander. Das Ähnlichkeitsmaß begünstigt dagegen die Assoziation mehrerer Zustandsmerkmale.*

gespeichert werden, so stimmen sie zwar in keinem Merkmal überein, werden aber dennoch schlecht bewertet: Sie stimmen darin überein, dass sie die anderen vier bekannten Merkmale nicht berücksichtigen. Dies wird durch (2.19) ebenfalls “bestraft”. Am besten werden Mengen von Assoziationen bewertet, in denen diese in möglichst wenigen Zustandsmerkmalen übereinstimmen, jedoch dennoch möglichst viele der bekannten Zustandsmerkmale in den einzelnen Assoziationen miteinander verknüpfen. Damit wirkt auch der simulierte Hippocampus wie sein natürliches Vorbild als Konfigurator weitläufiger Assoziationen.

### Übertragung ins EG

Für Experimente, in denen nur wenige Zustandsmerkmale eine Rolle spielen, wie das später in Abschnitt 4. präsentierte, werden aus den Assoziationen, die bereits an den Hippocampus übertragen wurden, alle möglichen Kombinationen gebildet und nach dem Ähnlichkeitsmaß (2.18) bewertet. Die Menge von Assoziationsmustern mit dem kleinsten Ähnlichkeitsmaß wird dann zur Integration ins Extensionsgedächtnis ausgewählt. Das Schwächerwerden von Gedächtnisspuren des Hippocampus, wie es in [10, McClelland et al.] vorausgesetzt wird, ist noch nicht berücksichtigt, ein einmal vom Objekterkennungssystem an den Hippocampus übertragenes Mu-

---

ster bleibt für immer dort gespeichert. Für Experimente mit vielen Zustandsmerkmalen muss das Bilden der Potenzmenge aller Assoziationen durch ein ökonomischeres Verfahren ersetzt werden. Dies ist technisch kein besonders großes Problem, die Bildung der Potenzmenge aller Assoziationen wurde vor allem zu Testzwecken gewählt, um sicher eine Teilmenge auswählen zu können, auf der das Ähnlichkeitsmaß ein Minimum hat.

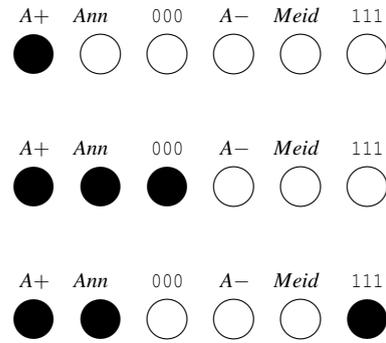
### 3.3 Das EG als Hopfield-Netz

Ein neuronales Netzwerkmodell wird nicht aus Gründen biologischer Plausibilität gewählt, um die assoziativen Fähigkeiten des Extensionsgedächtnisses zu modellieren. Obwohl neurobiologische Forschungen nahelegen, dass Prozesse des Gehirns durch künstliche neuronale Netzwerke gut simuliert werden können, ist der Abstraktionsgrad der meisten in der Praxis eingesetzten Modelle — so auch des hier verwendeten — sehr hoch. Außerdem stellt sich die Frage, warum psychologische Phänomene, die von neurobiologischen Prozessen zum Teil weit entfernt erscheinen, gerade durch die Modellierung einzelner Nervenzellen abgebildet werden sollten. Bestimmte kognitive Fähigkeiten wie z.B. das Ziehen logischer Schlussfolgerungen ließen sich mit Methoden der klassischen Künstlichen Intelligenz bereits befriedigend simulieren, der Einsatz neuronaler Netze erscheint hier unnötig. Die Entscheidung, das Extensionsgedächtnis durch ein neuronales Netzwerk zu modellieren, wird deshalb im Folgenden mit der Charakteristik der Informationsverarbeitung durch neuronale Netzwerkmodelle begründet.

#### Weitläufiges assoziatives Gedächtnis

Die Verknüpfung einzelner Wahrnehmungen zu weitläufigen assoziativen Netzwerken ist eines der wichtigsten Merkmale des Fühlens. Assoziativität kann durch ein Hopfield-Netz wie oben erläutert gut dargestellt werden. Als “weitläufig” kann man die Assoziationen innerhalb dieses Modells einerseits deswegen bezeichnen, da sie alle verfügbaren Statusinformationen (Affekte, Verhaltensweisen, Objektkategorien und der Befriedigung einzelner Bedürfnisse), die innerhalb des Modells vorkommen, miteinander verknüpfen.

Beziehungen zwischen den einzelnen Assoziationen werden durch die Hebb-Regel hergestellt: Die Summe über (2.8) legt fest, dass jedes ins Hopfield-Netz übertragene Assoziationsmuster gleichen Anteil an der Festlegung jedes einzelnen Verknüpfungsgewichtes hat. Damit wird auch zwischen den einzelnen Assoziationsmustern der größtmögliche Verknüpfungsgrad hergestellt. Die Stärke der Verknüpfungen zwischen zwei Nervenzellen eines Hopfield-Netzes kann als Stärke der Assoziation zwischen den entsprechenden Einzelmerkmalen interpretiert werden. Ein positiver Wert entspricht einer assoziativen Verknüpfung der beiden Merkmale in den meisten gespeicherten Mustern, ein negativer Wert tritt dann auf, wenn die beiden Merkmale in den meisten gespeicherten Mustern nicht assoziativ verknüpft sind. Dass der Wert einzelner Verknüpfungsgewichte jedoch nicht besonders aussagekräftig ist, zeigt das folgende Beispiel. Im Hopfield-Netz seien die drei Muster



gespeichert. Dies führt nach (2.8) zur Gewichtsmatrix

	A+	A-	Ann	000	100
A+	0.6	-0.6	0.2	-0.2	-0.2
A-	-0.6	0.6	-0.2	0.2	0.2
Ann	0.2	-0.2	0.6	0.2	0.2
000	-0.2	0.2	0.2	0.6	-0.2
100	-0.2	0.2	0.2	-0.2	0.6

Diese ist hier der Übersichtlichkeit halber als Verknüpfungstabelle dargestellt. Man findet den auf den ersten Blick merkwürdigen Effekt, dass positiver Affekt *negativ* mit der Objektkategorie 000 verknüpft ist (das entsprechende Gewicht hat den Wert -0.2). Dies kommt einfach dadurch zustande, dass positiver Affekt und 000 nur in einem der drei Muster zusammen auftreten, während in den anderen zwei Fällen positiver Affekt mit anderen Merkmalen verknüpft ist (siehe (2.8)). Dass im Netz sehr wohl eine Assoziation gespeichert ist, die positiven Affekt und 000 zusammen mit Annäherungsverhalten verknüpft, ist aus den einzelnen Verknüpfungsgewichten nicht ersichtlich.

### Implizites Gedächtnis

Um auf die im Netz gespeicherten Assoziationen schließen zu können, muss das Netz mit einem Vektor abgefragt werden, der den Zustand aller miteinander assoziierbarer Merkmale vorgibt. Das Ergebnis wird nach (2.6) unter paralleler Berücksichtigung *aller* im Hopfield-Netz vorhandenen Verknüpfungsgewichte berechnet. Dies zeigt, dass das Hopfield-Netz ein implizites Gedächtnis (siehe Fußnote 9, S.26) darstellt: Systeme wie Objekterkennungssystem und Intuitive Verhaltenssteuerung bekommen vom Extensionsgedächtnis nur eine "Antwort", von der sie aber nicht beurteilen können, wie sie zustande kam. In unserem Modell ergibt sich dies schon dadurch, dass andere Systeme nicht "beurteilen" können, was während möglicherweise mehrerer hintereinander ausgeführter Schaltschritte im Hopfield-Netz geschieht (siehe Abschnitt 2.3). Zur charakteristischen Arbeitsweise eines expliziten Gedächtnisses würde es dagegen passen, die Stärke von einzelnen Verknüpfungsgewichten als Hinweis dafür zu nehmen, welche Merkmale assoziiert sind und welche nicht. Auf diese Weise ist im Beispielnetz allerdings höchstens ablesbar, dass positiver und negativer Affekt gewöhnlich nicht zusammen auftreten, da hier das relativ stark negative Verknüpfungsgewicht -0.6 besteht.

---

### 3.4 Affekteinflüsse

Die in Abb. 1.2 dargestellte Informationsübertragung zwischen Objekterkennungssystem, Hippocampus und Extensionsgedächtnis findet nicht ständig statt, sondern wird durch die Affektdynamik (siehe Abschnitt 4.) beeinflusst. Nach der PSI-Theorie können Informationen nur dann ins Extensionsgedächtnis übertragen werden, wenn ein Wechsel von einer Aktivierung des Objekterkennungssystems zu einer Aktivierung des Extensionsgedächtnisses stattfindet. Deshalb werden Informationen aus dem Hippocampus nur dann ans Extensionsgedächtnis weitergegeben, wenn dieses aktiviert ist. Die Zustandsprotokolle des Objekterkennungssystems werden dagegen auch dann an den Hippocampus weitergeleitet, wenn das Objekterkennungssystem aufgrund der Affektdynamik deaktiviert ist. Dies kann zu Verzerrungen der Wahrnehmung führen (siehe Abschnitt 1.1). Bitstrings, die die einzelnen Objektkategorien bezeichnen, werden verändert “wahrgenommen”, z.B. wird statt eines Objektes der Kategorie 111 der um zwei Stellen verschobene Bitstring 100 ins Zustandsprotokoll aufgenommen. Dies kann dazu führen, dass ein eigentlich negativ besetztes als neutrales Objekt ins Extensionsgedächtnis übertragen wird: Nach Abschnitt 2.4 ist es möglich, dass die Kategorie 111 nicht so weit gefasst ist, dass sie Objekte mit dem Bitstring 100 noch enthält. Demnach würde ein Bedürfnis, das die Kategorie 111 als schädlich ansieht, das Objekt 100 nicht mit negativem Affekt bewerten. (siehe Abschnitt 2.4)

### 3.5 Verhaltenssteuerung durch das EG

In Abb. 1.2 ist nicht nur die Informationsübertragung zwischen Objekterkennung und Extensionsgedächtnis dargestellt, sondern es führt auch ein Pfeil vom Extensionsgedächtnis zur Verhaltenssteuerung. Hiermit wird angedeutet, dass die intuitive Verhaltenssteuerung Informationen des Extensionsgedächtnisses nutzen kann, um die Adäquatheit von Verhalten zu prüfen. Dies geschieht, indem Prämissen von Verhaltensprogrammen den Wert ihrer Aktivierung anhand von Abfragen des Extensionsgedächtnisses festlegen. Ein Beispiel wird in der Beschreibung des Simulationsexperimentes geliefert. Dort wird das Extensionsgedächtnis nach einer Verknüpfung der Verhaltensweise mit positivem Affekt befragt.

#### **Bereitstellung von Handlungsalternativen**

Wie in Abschnitt 1. dargestellt, besteht die Aufgabe des Extensionsgedächtnisses bei der Verhaltenssteuerung darin, möglichst viele ähnlich akzeptable Handlungsalternativen unter Berücksichtigung zahlreicher Kontextmerkmale der Verhaltenssteuerung gewissermaßen “vorschlagen”. Dies ist mit unserem Vorgehen gut abgebildet, da sich über unterschiedliche Prämissen (siehe Abschnitt 3.1) sehr verschiedene, sogar gegensätzliche Tendenzen des Verhaltens vorbereiten lassen, wobei es dann an der intuitiven Verhaltenssteuerung liegt, ein Verhaltensprogramm auszuwählen.

#### **Kongruenzorientierte Aufmerksamkeit**

Die hier verfolgte Strategie passt außerdem gut zum Konzept der *kongruenzorientierten Aufmerksamkeit* der PSI-Theorie (siehe Abschnitt 1.3). Durch Kopplung vieler Prämissen mit dem Extensionsgedächtnis ist es möglich, das charakteristische Warten auf eine seltene Gelegenheit (*Vigilanz*), die durch das Zusammenkommen vieler Situationsmerkmale nur sehr unscharf beschrieben werden kann, zu modellieren: Die “Kongruenz” der aktuellen Situation mit der

---

seltenen Gelegenheit für eine bestimmte Handlungsweise wird festgestellt, indem sehr viele Prämissen das Extensionsgedächtnisses daraufhin abfragen, ob die Bedingungen für eine bestimmte Handlungsweise erfüllt sind. Das komplexe Verhalten kann durch ein aus mehreren Basisverhaltensweisen zusammengesetztes Verhaltensprogramm (siehe Abschnitt 3.3) dargestellt werden, das bei hoher Aktivierung vieler einzelner Prämissen ausgelöst wird. Zu prüfen ist aber, ob sich der gewünschte Effekt nicht auch dann ergibt, wenn nur viele einzelne, einfachere Verhaltensprogramme starken Gebrauch vom assoziativen Wissen des Extensionsgedächtnisses machen.

### **Verhaltenssteuerung und der Kausalitätsbegriff Humes**

Hier stellt sich wiederum die Frage, ob die modellierte Person sich nicht ein wenig unkritisch auf ihr Erfahrungswissen verlässt. Wenn ein Verhaltensprogramm in seiner Prämisse das Extensionsgedächtnis befragt, ob das Ausführen des Verhaltens in der aktuellen Situation mit positivem Affekt assoziiert ist, wird ein in der Vergangenheit protokolliertes gleichzeitiges Auftreten von positivem Affekt in Zusammenhang mit einer bestimmten Verhaltensweise als Kausalbeziehung interpretiert. Damit handelt die simulierte Person ganz entsprechend dem *Kausalitätsbegriff* Humes: Eine Assoziation, deren Elemente nach dem Prinzip der *Berührung in Zeit und Raum* verknüpft worden sind, wird gewissermaßen automatisch als *Kausalbeziehung* interpretiert.<sup>16</sup> Sämtliche Kausalbeziehungen, die der Mensch zu erkennen meint, sind einzig durch *Gewöhnung* an das wiederholte zeitgleiche Auftreten zweier Dinge entstanden.<sup>17</sup> Damit löst Hume das Problem, dass bei ausschließlichem Rückgriff auf die Erfahrung, das Erkennen von Ursache-Wirkungs-Beziehungen schwer zu erklären ist — allerdings schlicht dadurch, dass er Kausalität grundsätzlich leugnet bzw. als Illusion abtut. Was würde wohl das Denken, jenes System der PSI-Theorie, dem man das Erkennen von Ursache-Wirkungs-Beziehungen zuschreiben muss, dazu sagen?

## **4. Experiment zu den Inhalten des EG**

In diesem Experiment wird gezeigt, dass sich allein durch die Parametrisierung mit den verschiedenen Affektsensibilitäten des STAR-Modells (siehe 4.4) inhaltliche Unterschiede bei den im Extensionsgedächtnis gespeicherten Assoziationen ergeben.

---

<sup>16</sup>Hume schreibt über die Art des Menschen, Ursache-Wirkungs-Beziehungen zu erkennen: “Nehmen wir weiter an, er (der Mensch) hätte mehr Erfahrung gewonnen und so lange in der Welt gelebt, um beobachtet zu haben, daß ihm geläufige Gegenstände oder Ereignisse konstant miteinander verbunden sind. Was ist die Folge dieser Erfahrung? Er leitet unmittelbar die Existenz des einen Gegenstandes aus der Erscheinung des anderen ab. Er hat jedoch — trotz seiner ganzen Erfahrung — keine Vorstellung oder Kenntnis der geheimen Kraft erlangt, durch die der eine Gegenstand den anderen hervorbringt, noch wird er durch irgendeinen Denkvorgang zu einer solchen Folgerung verpflichtet; und dennoch sieht er sich veranlaßt, so zu folgern.” [5, Hume, ]

<sup>17</sup>“Dieses Prinzip ist Gewohnheit (*custom*) oder herkömmliche Lebenspraxis (*habit*). Wo immer die Wiederholung einer bestimmten Handlung oder eines Vorganges das Verlangen hervorruft, dieselbe Handlung oder denselben Vorgang zu erneuern, ohne dazu durch einen Denktakt oder Verstandesvorgang gedrängt zu werden, sagen wir stets, dieses Verlangen sei die Wirkung der *Gewohnheit*. [...] Sicherlich nähern wir uns hier einem sehr einleuchtenden, wenn nicht wahren Satz, indem wir feststellen, daß wir gemäß einem konstanten Zusammenhang zweier Gegenstände — z.B. Hitze und Feuer, Gewicht und Masse — einzig durch Gewohnheit bestimmt werden, das eine beim Auftreten des anderen zu erwarten. Das scheint die einzige Hypothese zur Erklärung der Schwierigkeit zu sein, weshalb wir aus tausend Fällen etwas ableiten, was wir aus einem einzigen Falle, der sich doch in keiner Weise von ihnen unterscheidet, nicht ableiten können.” [5, Hume]

---

## 4.1 Versuchsbeschreibung

Die Person wird mit nur zwei Objekten konfrontiert, die sich in einer bestimmten Entfernung befinden sollen, die zu Beginn der Simulation zufällig bestimmt wird. Die Entfernung beträgt wenigstens fünf und höchstens fünfzehn Schritte. Eines der beiden Objekte soll bei der Person positiven Affekt hervorrufen, das andere negativen Affekt. Da beide Objekte in zufälliger Reihenfolge nacheinander als Ziele ausgewählt werden, kann das eine als “angenehmes” Ziel verstanden werden, dessen Ansteuern wenig Schwierigkeiten macht, das andere dagegen als “unangenehmes” Ziel. Das “unangenehme” der beiden Objekte ist mit einem Bitstring markiert, bei dem alle Bits gesetzt sind. Durch ungenaue Wahrnehmungen werden häufig die zwei letzten Bits des Bitstrings durch das Objekterkennungssystem auf 0 gesetzt (siehe Abschnitt 1.1). Dies wird im weiteren Verlauf durch den verkürzten Bitstring 100 dargestellt. Wird das Objekt durch das Objekterkennungssystem genau wahrgenommen, sind also alle Bits gesetzt, wird dies durch den Bitstring 111 symbolisiert. Das “angenehme” Objekt erhält den Bitstring 000. Das Verhaltens-



Abbildung 4.14: Die beiden Objekte des Simulationsexperimentes tragen jeweils einen Bitstring und eine Distanz, die die Entfernung zur Person in Schritten angibt.

repertoire besteht aus nur einer Verhaltensroutine, der Annäherung (*Ann*) an das Zielobjekt. Das Verhalten wird ausgelöst durch die zwei Prämissen, *GetPosAff* und *MindNegAff*. *GetPosAff* und *MindNegAff* sind durch logisches ODER verknüpft, d.h. das Annäherungsverhalten wird dann aktiviert, wenn die Person “glaubt”, dass das anzusteuern Zielobjekt positiven Affekt auslöst oder negativen Affekt meidet. Dies wird durch eine Anfrage an das Extensionsgedächtnis-



Abbildung 4.15: Annäherungsverhalten wird nach Prüfung der Prämissen *GetPosAff* und *MindNegAff* ausgelöst: Die beiden Prämissen werden jeweils aktiv, wenn Annäherung an das aktuelle Zielobjekt im Extensionsgedächtnis mit positivem bzw. nicht mit negativem Affekt assoziiert ist.

nis festgestellt: *GetPosAff* erhält die Aktivierung 1 (vollständig aktiviert), wenn eine Anfrage an das Extensionsgedächtnis mit dem ausgewählten Zielobjekt und Annäherungsverhalten ein Assoziationsmuster zurückliefert, bei dem positiver Affekt aktiviert ist. Ist das Annäherungsverhalten dem Extensionsgedächtnis bisher nicht bekannt, da noch keine Assoziationen übertragen wurden, in denen Annäherungsverhalten vorkam, wird nur abgefragt, ob das Zielobjekt mit positivem Affekt assoziiert ist. Bestehen keine Assoziationen zu positivem Affekt, erhält die Prämisse die Aktivierung 0 (vollständig deaktiviert). Analog wird *MindNegAff* dann aktiviert, wenn Annäherungsverhalten und Zielobjekt (bzw. nur das Zielobjekt bei nicht bekanntem Annäherungsverhalten) *nicht* mit negativem Affekt assoziiert sind. Liefert eine entsprechende Abfrage des Extensionsgedächtnisses keinen negativen Affekt zurück, erhält diese Prämisse die Aktivierung 1, sonst 0. Im Gegensatz zur Prämisse *MinDist* (siehe 3.2) werden *GetPosAff* und

---

MindNegAff “scharf” aktiviert. Dies bedeutet, dass die Prämissen nur die “scharfen” Aktivierungen 0 und 1 haben können. Eine Annäherung an ein Zielobjekt ist nur möglich, solange die Person einen Abstand größer 0 hat. Das Experiment endet dann, wenn beide Zielobjekte erreicht worden sind — oder klar ist, dass nicht beide Objekte erreicht werden können: Dies geschieht dann, wenn beide Prämissen des Annäherungsverhalten deaktiviert sind, da das Zielobjekt mit negativem Objekt assoziiert ist. Hier endet das Experiment, sobald die Person keine neuen Informationen mehr ins Extensionsgedächtnis übertragen kann, d.h. wenn beide Zielobjekte dem EG bekannt sind. Das Experiment wird für alle Parametrisierungen des STAR-Modells (siehe Tabelle 2.1) für die Seeds 0s-26s des Quicksilver-internen Zufallsgenerators durchgeführt (siehe [6]).

## 4.2 Beobachtungen

Bei allen acht Parametrisierungen ergeben sich stets Paare oder Tripel, in Einzelfällen eine Vierer-Kombination von im Extensionsgedächtnis gespeicherten Assoziationen. Diese verändern sich nach einiger Zeit nicht mehr. Die jeweils zum Ende der Simulationsläufe erreichten stabilen Assoziationen sind in den Tabellen A.5, A.6, A.7 und A.8 dargestellt. Die Ergebnisse wurden nach Ähnlichkeit der stabilen Assoziationen in vier Gruppen eingeordnet (siehe Abb. 4.16). Diese Gruppen können anhand der Ausprägtheit ihres *Belohnungs-* bzw. *Bestrafungssysteme* unterschieden werden: Das Belohnungssystem ist für die Generierung positiver, das Bestrafungssystem für die Erzeugung negativer Affekte verantwortlich. *Belohnungssensitiv* werden die Persönlichkeitstypen genannt, die über ein mittel bis stark ausgeprägtes Belohnungssystem und ein schwach ausgeprägtes Bestrafungssystem verfügen — hierzu gehören also der *liebenswürdige*, der *ehrgeizige* und der *selbstbestimmte* Typ. Als *bestrafungssensitiv* werden analog die Typen bezeichnet, die ein mittleres bis starkes Bestrafungssystem und ein niedrig ausgeprägtes Belohnungssystem haben — dies sind die *sorgfältige*, die *selbstkritische* und die *zurückhaltende* Persönlichkeit. Die *eigenwillige* und die *ahnungsvolle* Persönlichkeit können in keine der beiden Gruppen eingeordnet werden und werden deshalb gesondert behandelt.

### Inhaltliche Unterschiede

Bei den Ergebnissen ist ein Übergang von den belohnungssensitiven Typen über die eigenwillige Persönlichkeit zu den bestrafungssensitiven Persönlichkeitstypen festzustellen: Die *belohnungssensitiven* Persönlichkeiten bilden *ausschließlich* Assoziationen, die positiven Affekt enthalten. Demgegenüber gibt es bei den *bestrafungssensitiven* Persönlichkeitstypen eine deutliche Menge von Assoziationen mit negativem Affekt. Zwischen diesen beiden Extremen steht die *eigenwillige* Persönlichkeit: Sie überträgt etwa gleich viele mit positivem Affekt verknüpfte Assoziationen ins Extensionsgedächtnis wie solche, die mit keiner Affektsorte verbunden sind, jedoch keine Assoziationen mit negativem Affekt. Eine Sonderstellung nimmt die *ahnungsvolle* Persönlichkeit ein: Dadurch, dass sie beide Affektsorten heraufreguliert, gelingt es ihr, typische Assoziationen sowohl der *belohnungssensitiven* als auch der *bestrafungssensitiven* Persönlichkeitstypen zu bilden.

### Leistungsunterschiede

Definiert man “Erfolg” dadurch, dass beide Objekte, d.h. sowohl das bedürfnisbefriedigende als auch das bedürfnisschädigende in möglichst kurzer Zeit erreicht werden sollen, ergibt sich

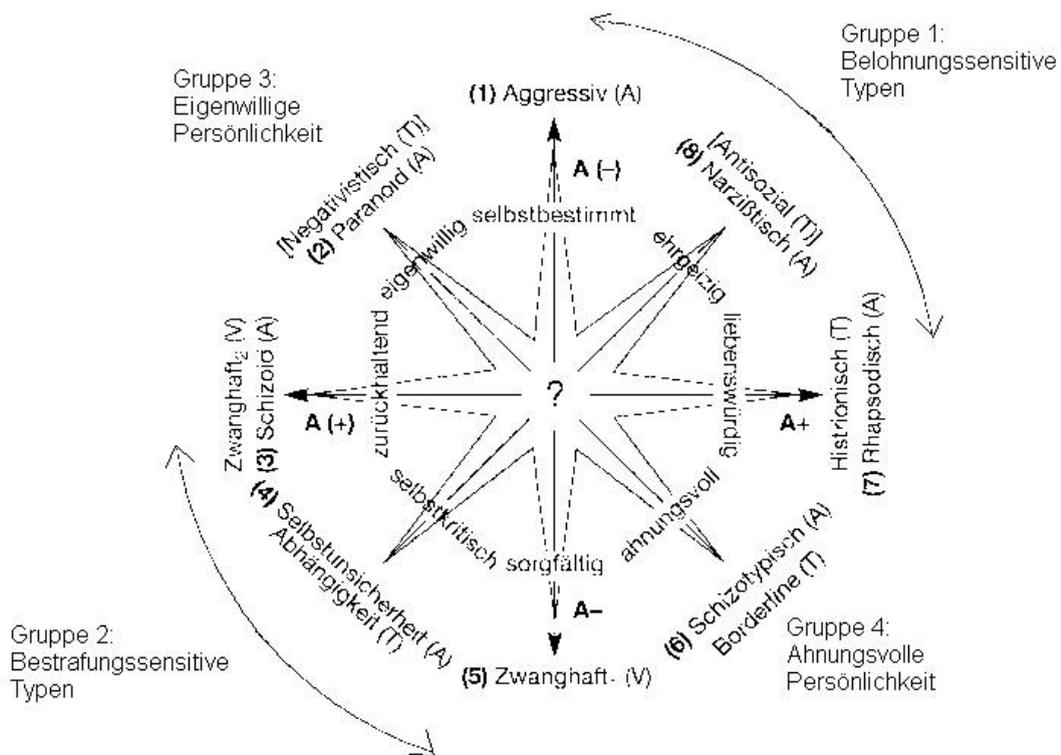


Abbildung 4.16: Die Persönlichkeiten des STAR-Modells können aufgrund ihres Verhaltens in einem Experiment zu inhaltlichen Unterschieden der im EG gespeicherten Assoziationsmuster in vier Gruppen aufgeteilt werden.

---

auch hier ein klares Bild: Erwartungsgemäß schneiden die *belohnungssensitiven* Typen am besten ab. Sie erreichen in allen 27 Läufen beide Zielobjekte. Der *ehrgeizige* Persönlichkeitstyp benötigt durchschnittlich deutlich weniger Zeitschritte als die *liebenswürdige* und die *selbstbestimmte* Persönlichkeit, die in allen Läufen in exakt der gleichen Anzahl von Schritten ans Ziel kommen. *Eigenwillige* benötigen gegenüber letzteren etwas mehr Schritte. Deutlich fällt der Unterschied zu den *bestrafungssensitiven* Typen aus: Hier werden nur in etwa der Hälfte der Fälle beide Objekte erreicht. Bildet man die durchschnittliche Anzahl von Schritten für die Läufe, in denen das Ziel erreicht wurde, ergibt sich für die *sorgfältige* und die *zurückhaltende* Persönlichkeit ein Wert, der etwa im Bereich der *liebenswürdigen* und der *selbstbestimmten* Persönlichkeit liegt. Berechnet man für diese allerdings den Durchschnitt nur auf den von den *bestrafungssensitiven* Typen erfolgreich abgeschlossenen Läufen, sind die Leistungen der *belohnungssensitiven* Persönlichkeitstypen deutlich besser. Dies kann so interpretiert werden, dass die *bestrafungssensitiven* Persönlichkeiten dann beide Ziele erreichen, wenn diese allen anderen Typen vergleichsweise “leicht” fallen, d.h. in relativ wenig Zeitschritten angesteuert werden. *Ahnungsvolle* Persönlichkeiten erreichen ebenfalls “leichte” Ziele, für diese benötigen sie wegen ihrer hohen Belohnungssensitivität sehr wenig Zeitschritte.

### 4.3 Auswertung

Die im Experiment gestellte Aufgabe besteht darin, ein “angenehmes” und ein “unangenehmes” Ziel in möglichst kurzer Zeit zu erreichen. Das jeweils aktuelle Ziel wird zufällig bestimmt, kann also nicht “selbstgesteuert” von der Person gewählt werden; d.h. das Selbstsystem kann keinen Einfluss auf die Auswahl nehmen. “Selbstgesteuert” ist dagegen die affektive Bewertung der assoziativen Verknüpfungen im Extensionsgedächtnis. Da das Extensionsgedächtnis zu Beginn eine “tabula rasa” ohne frühere Erfahrungen ist, entspricht die Situation des Experiments also dem Sammeln erster Erfahrungen in einem für die Person völlig unbekanntem Bereich, wobei die zu erreichenden Ziele von außen vorgegeben werden — dies ist vergleichbar mit den ersten Schultagen oder dem Beginn der Arbeit in einem neuen Beruf. Insgesamt wird das Experiment deshalb als Test für das Erkennen positiver Anreize in einer neuen Situation interpretiert. In diesem Abschnitt wird zunächst gezeigt, dass sich im Experiment — in Übereinstimmung mit den Annahmen der PSI-Theorie — im Extensionsgedächtnis ein positiver Bias ergibt. Weiterhin wird erläutert, warum sich die im Extensionsgedächtnis gespeicherten Assoziationen sich nach einer bestimmten Laufzeit nicht mehr verändern. Abschließend hauptanliegen dieses Abschnittes ist es, zu zeigen, dass sich sämtliche hier dargestellten Effekte allein durch den in Abschnitt 1.5 dargestellten Informationsübertragungsmechanismus vom Objekterkennungssystem über den Hippocampus zum Extensionsgedächtnis ergeben. Weitergehende Interpretationen des Experiments, die nur durch zusätzliche theoretische Überlegungen gemacht werden können, werden im abschließenden Abschnitt 4.4 behandelt.

#### Positiver Bias der Inhalte des EG

Es zeigt sich, dass das Extensionsgedächtnis tatsächlich den postulierten *positiven Bias*<sup>18</sup> hat — siehe für diesen Abschnitt jeweils die Tabellen A.1-A.4 im Anhang. Die Mehrzahl der im Ex-

---

<sup>18</sup>“Normalerweise hat das integrierte Selbstsystem jedoch einen eingebauten *positiven Bias*: Die mit dem Extensionsgedächtnis verbundene kongruenzsensitive Aufmerksamkeitsform verstärkt Erfahrungen, die mit den eigenen Bedürfnissen, Wünschen und Erwartungen übereinstimmen. Erfolge werden eher wahrgenommen und ins Selbstsystem integriert als Misserfolge.”[9, Kuhl, S.171]

---

tensionsgedächtnis gespeicherten Assoziationen enthält positiven Affekt oder zumindest keine Verknüpfung mit positivem oder negativem Affekt. Extrem ist das Ergebnis der *belohnungs-sensitiven* Persönlichkeitstypen: In dieser Gruppe sind *sämtliche* Assoziationen mit positivem Affekt verknüpft.

Nur bei *bestrafungssensitiven* Typen überwiegt teilweise negativer Affekt gegenüber positivem Affekt: Während bei der *sorgfältigen Persönlichkeit* gleich viele Assoziationen positiven wie negativen Affekt enthalten (jeweils 16), ergibt sich beim *selbstkritischen Typ* ein leichtes, beim *zurückhaltenden* ein relativ starkes Übergewicht an Assoziationen mit negativem Affekt: 13 Assoziationen enthalten jeweils positiven Affekt, gegenüber 17 negativ besetzten bei der *sorgfältigen* und 21 mit negativem Affekt verbundenen Assoziationen bei der *zurückhaltenden* Persönlichkeit. Insgesamt finden sich bei den *bestrafungssensitiven* Typen jedoch vor allem Assoziationen ohne Verbindung zu einer der beiden Affektsorten.

Deutlich zeigt sich der positive Bias auch beim *eigenwilligen* und beim *ahnungsvollen Typ*: Obwohl *eigenwillige* Persönlichkeiten sowohl positiven als auch negativen Affekt herabregulieren, übertragen sie *keine einzige* Assoziation mit negativem Affekt ins Extensionsgedächtnis, dagegen ist knapp die Hälfte der Assoziationen mit negativem Affekt verbunden; die meisten sind ohne Affektbesetzung.

Beim *ahnungsvollen Typ* liegt die entgegengesetzte Situation vor; dieser Typ ist durch Sensibilität für beide Affektsorten charakterisiert. Dennoch werden Assoziationen mit positivem Affekt gegenüber mit negativem Affekt assoziierten Inhalten bevorzugt: 30 Assoziationen enthalten positiven Affekt und sieben weniger negativen Affekt. Wegen der hohen Sensibilität für beide Affektsorten kommt nur eine Assoziation ohne Affektverknüpfung vor. Der positive Bias lässt sich also in allen vier Gruppen deutlich erkennen.

Ein weiterer Beleg für die positive Verzerrung der Inhalte des Extensionsgedächtnis ist an der Repräsentation der Objektkategorie 111 erkennbar: Nur den *bestrafungssensiblen* Typen gelingt es — äußerst selten —, tatsächlich die Kategorie 111 ins Extensionsgedächtnis zu übertragen. In den meisten Fällen wird diese Kategorie als 100 im Extensionsgedächtnis abgelegt. Dies kommt durch den Einfluss positiven Affekts auf die Objekterkennung zustande: Positiver Affekt führt zu einer Deaktivierung des Objekterkennungssystems, was in unserem Modell die in Abschnitt 1.1 erläuterte Veränderung der Wahrnehmung zur Folge hat. Auch dieser Effekt kann als Ausdruck des positiven Bias des Extensionsgedächtnis gewertet werden: Die für die Bewertung der Objekte zuständigen Bedürfnisse bringen die Objektkategorie 100 möglicherweise gar nicht mit negativem Affekt in Verbindung, da in erster Linie die durch den Bitstring 111 gekennzeichneten Objekte als “gefährlich” angesehen werden und zur Generierung von negativem Affekt führen. Dies reduziert für die Person die Gefahr, sich durch Abruf von Assoziationen negativem Affekt auszusetzen: Führt eine Abfrage des Extensionsgedächtnisses zu einer Assoziation mit der Kategorie 111, wird mit größerer Wahrscheinlichkeit negativer Affekt ausgelöst, als durch die Objektkategorie 100. So setzt das Extensionsgedächtnis die “Bedrohlichkeit” seiner Inhalte herab — stößt eine Person auf die Kategorie 100, kann sie erkennen, dass es sich um eine potentielle Gefahr handelt, ohne jedoch gleich durch negativen Affekt gelähmt zu werden.

Zum Abschluss dieses Abschnitts muss noch darauf eingegangen werden, wodurch der positive Bias des Extensionsgedächtnisses in diesem Modell zurückzuführen ist. Die Ursache ist einzig und allein in der Art und Weise der Informationsübertragung vom Hippocampus zum Extensionsgedächtnis zu suchen: Um die Empfindlichkeit des Hippocampus gegenüber negativem Affekt zu berücksichtigen, werden nur dann Assoziationen übertragen, wenn das Extensionsgedächtnis durch entspannte Affektlage aktiviert ist. Liegt also negativer Affekt vor, wird dies zwar vom Objekterkennungssystem in einem Zustandsprotokoll festgehalten; allerdings ist es

---

unwahrscheinlich, dass eine Übertragung ins Extensionsgedächtnis stattfindet, da dieses durch negativen Affekt deaktiviert wird.

### **Stabile Assoziationen**

Es erscheint überraschend, dass die Inhalte des Extensionsgedächtnisse sich nach einer bestimmten Anzahl von Schritten nicht mehr verändern. Eine “Abbruchbedingung” für Veränderungen des Extensionsgedächtnisses ist im Modell schließlich nicht vorgesehen. Das Verhalten des Modells wird jedoch einleuchtend, wenn man beachtet, dass die Informationsübertragung ins Extensionsgedächtnis durch die Minimierung des Ähnlichkeitsmaßes (2.19) gesteuert wird: Dieses ist nach unten durch -1 beschränkt, spätestens bei Erreichen dieses Wertes, kann sich das Extensionsgedächtnis nicht mehr verändern.

Die Bildung stabiler Assoziationen ist durchaus realistisch: Ein wohlbekanntes Merkmal des Alterns ist die abnehmende Empfänglichkeit für neues Erfahrungswissen. Die Unfähigkeit des zugrundeliegenden neuronalen Netzes, sich durch die Integration neuer Erfahrungen noch zu verändern, nachdem es bereits einen relativ stabilen Zustand erreicht hat, wird auch in [10, McClelland et al., S.448] in Erwägung gezogen, um diesen Alterungseffekt zu erklären. Für unser Experiment genügt schon eine bescheidenere Interpretation: Soll — wie zu Beginn vorgeschlagen — das Sammeln neuer Erfahrungen in einem bisher unbekanntem Bereich simuliert werden, erinnert das Modellverhalten der Gewöhnung an eine neue Situation: Nach einer Anfangsphase, in der viel Wissen verarbeitet und erste Erfahrungen häufig korrigiert werden müssen, wird ein Zustand erreicht, in dem eine ständige Anpassung des Erfahrungswissens nicht mehr notwendig ist.

### **Die Assoziationen $A--(000)$ und $A+-Ann-(100)$**

Einige Assoziationen, die sich im Experiment ergeben, erscheinen paradox, da sie nicht etwa die “Wirklichkeit” der im Modell simulierten “Welt” abbilden — sondern bei *bestrafungssensitiven* und *ahnungsvollen* Typen in einigen Fällen sogar das Gegenteil: Wie ist es zu erklären, dass die Assoziationen  $A--(000)$  und  $A+-Ann-(100)$  im Extensionsgedächtnis gespeichert werden können?

Warum diese Kombination von Assoziationen mit recht hoher Wahrscheinlichkeit ins Extensionsgedächtnis übertragen wird, ist relativ leicht zu erklären: Interessanterweise taucht die Assoziation  $A+-Ann-(100)$  in allen Gruppen auf — außer bei *eigenwilligen* Persönlichkeiten. Zunächst ist klar, dass es nicht einfach ist, eine Repräsentation der Objektkategorie 111 zu bilden — wie schon oben erwähnt, ergibt sich der Bitstring 100 durch eine Ungenauigkeit der Wahrnehmung bei deaktiviertem Objekterkennungssystem (siehe Abschnitt 1.1). Die Schwierigkeit besteht darin, dass zum Informationsaustausch von Objekterkennungssystem und Extensionsgedächtnis eine gewisse Menge positiven Affekts notwendig ist und das betreffende Objekt ja gerade negativen Affekt hervorruft. Eine gute Chance, es im Extensionsgedächtnis zu speichern, besteht dann, wenn die Person sich im vorangehenden Zeitschritt an das Objekt angenähert hat: In diesem Fall muss positiver Affekt vorhanden gewesen sein, da die Aktivierung der Verhaltenssteuerung, die für das Ausführen von Verhalten nötig ist, positiven Affekt erfordert. Ein Blick auf sämtliche gespeicherten Muster ergibt, dass (100) tatsächlich ausschließlich zusammen mit *Ann* gespeichert wird. Die Kombination der beiden Assoziationen  $A--(000)$  und  $A+-Ann-(100)$  ist nun leicht erklärbar:  $A--(000)$  ist das Muster, das zu  $A+-Ann-(100)$  den maximalen “Overlap” (2.15) hat. Gelangt diese Assoziation also jemals in den Hippocampus,

---

ergibt sich aus dem Ähnlichkeitsmaß zwangsläufig, dass diese Kombination ins Extensionsgedächtnis übertragen wird.

Wie empfindlich der Mechanismus selbst auf kleine Änderungen reagiert, zeigt sich, wenn man betrachtet, was geschieht, wenn *Ann*-(100) ohne affektive Bewertung gespeichert wird: In diesem Fall behält das System seine "Freiheit" sowohl eine Assoziation der "angenehmen" Objektkategorie (000) mit positivem oder negativem Affekt zu bilden. Dass diese "Freiheit" durchaus genutzt wird, wird später in der Diskussion der Inhalte des Extensionsgedächtnisses gezeigt werden.

An diesem Beispiel ist ersichtlich, dass die Inhalte des Extensionsgedächtnisses nicht unbedingt die Realität abbilden müssen. Stattdessen werden insbesondere Kategorien, die mit negativem Affekt besetzt sind, "umbewertet".

### Inhaltlicher Vergleich der Assoziationen der einzelnen Gruppen

Die inhaltlichen Unterschiede können auf den minimalen Wert des Overlaps (2.15) der im Extensionsgedächtnis gespeicherten Assoziationen zurückgeführt werden. Im Experiment tauchen fünf Zustandsmerkmale auf; dies sind: Positiver und negativer Affekt ( $A+$  und  $A-$ ), Annäherungsverhalten (*Ann*) sowie die Objektkategorien 000 und 100 (100 ist eine Repräsentation der Kategorie 111, siehe den vorhergehenden Abschnitt). Als minimaler Wert des Overlaps (2.15) ergibt sich bei fünf Zustandsmerkmalen -5, dieses Minimum wird dann erreicht, wenn zwei Assoziationen auf keinem Merkmal übereinstimmen. Nicht alle Persönlichkeitstypen sind jedoch in der Lage, eine Kombination von Assoziationen mit minimalem Overlap zu bilden. In diesem Fall werden Kombinationen von Assoziationen mit dem nächstniedrigeren Overlap im Extensionsgedächtnis gespeichert.

**Belohnungssensitive Persönlichkeitstypen:** Die *belohnungssensitiven* Persönlichkeiten verknüpfen alle Assoziationen mit positivem Affekt und übertragen keinen negativen Affekt ins Extensionsgedächtnis (siehe Tabelle A.1). Weil die Objektkategorie (100) nur in Verbindung mit Annäherungsverhalten (*Ann*) im Extensionsgedächtnis gespeichert wird, gibt es für *belohnungssensitive* Typen nur eine Kombination von Assoziationen (siehe Tabelle A.5) mit dem minimalen Overlap (2.15) von -1 (siehe Abb. 4.17). Diese wird zwangsläufig dann ins Extensionsgedächtnis übertragen, sobald die beiden Assoziationen gelernt worden sind. Die Assoziationen

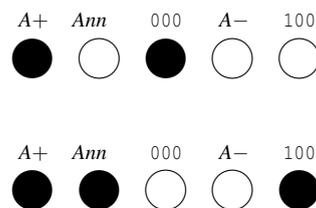


Abbildung 4.17: Diese Kombination von Assoziationen hat für den Overlap (2.15) den Wert -1. Dies ist der minimale Overlap, den belohnungssensitive Persönlichkeiten im Experiment erreichen können, da sie alle Assoziationen mit positivem Affekt verknüpfen.

$A+$ -(000) und  $A+$ -*Ann*-(100) bedeuten, dass *belohnungssensitive* Persönlichkeiten den Anreiz

der “angenehmen” Objektkategorie 000 “erkennen” und das “unangenehme” Objekt “umbewerten”. Im Experiment haben sie somit beste Voraussetzungen, beide Zielobjekte schnell zu erreichen. Dies wird später mit der *Selbstmotivierung* in Zusammenhang gebracht. Die anderen drei Kombinationen, die sehr selten ins Extensionsgedächtnis übertragen werden, haben jeweils einen Overlap (2.15) von 1; in diesen Fällen ist entweder die Assoziation A+-(000) oder A+-Ann-(100) vor Erreichen beider Zielobjekte noch nicht im Hippocampus gebildet worden.

Die Belohnungssensitivität zeigt sich darin, dass positiver Affekt sogar isoliert ins Extensionsgedächtnis übertragen wird. Dies ist bei den anderen Persönlichkeitstypen — außer der *ahnungsvollen* Persönlichkeit nicht möglich, da es ihnen nicht gelingt, positiven Affekt ohne das Vorhandensein von Anreizobjekten auf einem erhöhten Niveau zu halten. Selbstverständlich ist das Lernen der “Assoziation” A+ nicht besonders sinnvoll, damit wird es unmöglich, dass Extensionsgedächtnis mit dem Muster A+ zu befragen — es würde nur die gespeicherte Assoziation A+ zurückgeben. Dies lässt darauf schließen, dass es *belohnungssensitiven* Persönlichkeit gelegentlich schwer fallen könnte, durch Abfrage des Extensionsgedächtnisses die *Ursachen* von positivem Affekt anzugeben — dies ergibt sich daraus, dass sie positiven Affekt in manchen Situationen eher undifferenziert erleben, da schon ein paar Zeitschritte zurückliegende Störungen durch positiven Affekt nicht herabreguliert werden.

**Eigenwillige Persönlichkeit:** Noch eintöniger sind die im Extensionsgedächtnis repräsentierten Inhalte bei der *eigenwilligen* Persönlichkeit (siehe Tabelle A.3). Wie die *belohnungssensitiven* Persönlichkeitstypen ist sie nicht in der Lage, negativen Affekt ins Extensionsgedächtnis zu übertragen. Hinzu kommt, dass es ihr nicht gelingt, Ann-(100) mit positivem Affekt zu verknüpfen. Auf den ersten Blick scheint das damit zusammenzuhängen, dass sie positiven Affekt dämpft — doch warum gelingt es den *Eigenwilligen* dann, immerhin die Objektkategorie 000 assoziiert mit positivem Affekt ins Extensionsgedächtnis zu übertragen? Die Kombination A+-(000) und Ann-(100) ergibt sich deswegen so häufig — in 24 von 27 Läufen (siehe Tabelle A.7) —, weil diese mit dem Wert -3 einen geringeren Overlap (2.15) hat als A+-Ann-(100) und A+-(000), siehe Abb. 4.18. Selbst wenn also auch die Assoziation A+-Ann-(100) in den

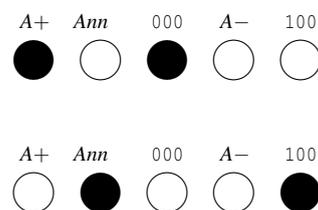


Abbildung 4.18: Der *eigenwilligen* Persönlichkeit gelingt es, Annäherungsverhalten und die Objektkategorie 100 auch ohne Verknüpfung mit positivem Affekt miteinander zu assoziieren. Dadurch wird die hier abgebildete Assoziation statt der in Abb. 4.17 gezeigten ins Extensionsgedächtnis übertragen.

Hippocampus übertragen wird, kann sie sich gegenüber Ann-(100) nicht durchsetzen. Die übrigen drei Fälle lassen sich wie bei den *belohnungssensitiven* Typen darauf zurückführen, dass es nicht gelang, eine der beiden Assoziationen A+-(000) bzw. Ann-(100) zu erlernen. Die dominante Kombination A+-(000) und Ann-(100) unterscheidet sich von den *belohnungssensitiven*

---

Persönlichkeiten dadurch, dass zwar der positive Anreiz der Objektkategorie 000 erkannt wird, jedoch die Annäherung an das “unangenehme” Ziel nicht positiv “umbewertet” wird (siehe den Abschnitt über *Selbstmotivierung*).

**Bestrafungssensitive Persönlichkeitstypen:** Eine größere Vielzahl unterschiedlicher ins Extensionsgedächtnis übertragener Informationen zeigt sich bei den *bestrafungssensitiven* Persönlichkeitstypen (siehe Tabelle A.2). Dies ergibt sich daraus, dass hier die ganze Bandbreite affektiver Bewertungen vertreten ist. Die *bestrafungssensiblen* Typen sind die einzige Gruppe, der es gelingt, die Objektkategorie 111 tatsächlich ins Extensionsgedächtnis zu übertragen — und diese nicht durch die bei deaktiviertem Objekterkennungssystem wahrgenommene Kategorie 100 zu repräsentieren. Die sich ergebende Kombination von  $A--Ann-(111)$  und  $A+- (000)$  (siehe Tabelle A.6) ist ein Beispiel dafür, dass eine Menge von Assoziationen, auch wenn sie ein hohes Ähnlichkeitsmaß (2.19) hat, nicht zwangsläufig oft ins Extensionsgedächtnis übertragen werden muss. Dass diese beiden Kombinationen trotz des minimalen Overlaps (2.15) von -5 nicht häufiger ins Extensionsgedächtnis übertragen wird, liegt daran, dass auch *bestrafungssensitiven* Persönlichkeiten die Kategorie 111 nur selten ins Extensionsgedächtnis übertragen können: Meistens übertragen auch sie die Assoziation  $Ann-(100)$ . Die Kombination der Assoziationen  $A--Ann-(111)$  und  $A+- (000)$  ist im Vergleich zu anderen stabilen Assoziationsmustern eine ziemlich “wirklichkeitsnahe” Repräsentation der Situation des Experiments: Die Objektkategorie 000 ist mit positivem Affekt besetzt und Annäherung an Objekte der Kategorie 111 führt zu negativem Affekt. Doch dieser “Realismus” zahlt sich, wie auch Kuhl in [9] bemerkt, nicht aus: Tatsächlich haben Depressive ein sehr viel realistisches Bild ihrer Handlungsmöglichkeiten als Nicht-Depressive — für das Meistern schwieriger Situationen ist dieser Realismus allerdings nicht geeignet.<sup>19</sup> Dies zeigt sich auch an den Ergebnissen des Experimentes: Assoziative Verknüpfungen mit negativem Affekt führen dazu, dass in vielen Fällen nicht beide Objekte erreicht werden können, da die beiden Prämissen *GetPosAff* und *MindNegAff* nicht erfüllt sind und das Ausführen des Annäherungsverhaltens verhindert wird. Die auf den ersten Blick überraschende Tatsache, dass die Assoziation  $A--(000)$  zwar in vielen Fällen, jedoch nicht zwangsläufig dazu führt, dass nicht beide Objekte angesteuert werden können, ist dadurch zu erklären, dass die Assoziation erst dann ins Extensionsgedächtnis übertragen wurde, *nachdem* das Objekt der Kategorie 000 bereits erreicht worden war.

Betrachten wir hierzu noch die zwei häufig vorkommenden Kombinationen,  $A+- (000)$  und  $Ann-(100)$  sowie  $A--(000)$  und  $Ann-(100)$ . Hier besteht im Gegensatz zu den vorigen Gruppen die “freie Wahl”, die Objektkategorie 000 mit positivem bzw. negativem Affekt zu verknüpfen. Tatsächlich wird diese “Freiheit” beim *sorgfältigen* Typ genutzt: Jeweils elfmal wird 000 mit  $A+$  bzw.  $A-$  verknüpft, beim *selbstkritischen* und beim *zurückhaltenden* Typ ergibt sich sogar ein Übergewicht der negativen Bewertung der Objektkategorie 000. Beim *zurückhaltenden* Typ kommt dies durch die starke Herabregulierung positiven Affekts zustande — ihm gelingt es nur sechsmal, die Kombination  $A+- (000)$  und  $Ann-(100)$  zu bilden, den anderen beiden *bestrafungssensiblen* Typen fast doppelt so häufig. Sowohl der *zurückhaltende* als auch der *selbstkritische* Typ bringen die Objektkategorie 000 etwas häufiger mit negativem Affekt in Verbindung als die *sorgfältige* Persönlichkeit. Für den *zurückhaltenden* Persönlichkeitstyp führt dies dazu,

---

<sup>19</sup>“Depressivität ist oft gerade mit einer *Reduktion* der für Nicht-Depressive typischen und bei positiver Grundstimmung zu erwartenden Überschätzung der realen Kontrollmöglichkeiten verknüpft (Alloy & Abramson, 1979). Depressive sehen die Realität demnach oft nicht pessimistisch, sondern allzu realistisch, d.h. ihnen fehlt die positive Verzerrung der Wirklichkeit, ohne die Nicht-Depressive in vielen schwierigen Lebenslagen nicht die Kraft hätten, die nötige Ausdauer zu zeigen.” [9, Kuhl, S.434]

---

dass er nur in 12 Läufen beide Ziele erreichen kann — dies ist das zweit schlechteste Ergebnis aller Parametrisierungen, nur *Ahnungsvolle* schneiden noch schlechter ab.

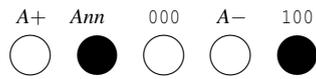
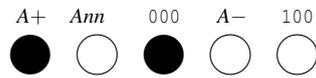
**Ahnungsvolle Persönlichkeit** Der *ahnungsvolle* Persönlichkeitsstil liefert im hier vorgestellten Experiment besonders interessante Ergebnisse (siehe Tabelle A.4). Durch die Sensibilität für positiven *und* negativen Affekt finden sich im Extensionsgedächtnis von Personen dieses Persönlichkeitstyps charakteristische Assoziationen sowohl der *belohnungssensitiven* als auch der *bestrafungssensitiven* Persönlichkeiten wieder (siehe Tabelle A.8): Die beiden Assoziationen  $A--(000)$  und  $A+-App-(000)$  stammen aus dem Repertoire der *belohnungssensitiven* Persönlichkeiten,  $A--(000)$  und  $A+-Ann-(100)$  kommt bei den *bestrafungssensitiven* Typen vor, typischer ist jedoch die Kombination  $Ann-(000)$  ohne affektive Bewertung. Durch die Sensibilität für beide Affektsorten kommt die Kombination  $A+-(000)$  und  $A+-App-(000)$  bei der *ahnungsvollen* Persönlichkeit besonders häufig vor (siehe Abb. 4.20 und Abschnitt “Die Assoziationen  $A--(000)$  und  $A+-Ann-(100)$ ”). Dies führt letztendlich zum extrem schlechten Ergebnis der *ahnungsvollen* Persönlichkeit, die nur in zehn Läufen beide Ziele erreicht. Das hier dargestellte fatale Zusammenwirken sowohl der Affektsensibilitäten und der Informationsübertragung durch den Hippocampus wirft die Frage auf, inwieweit das Modell sich an dieser Stelle noch realistisch verhält: Dass das Extensionsgedächtnis von *Ahnungsvollen* durch die Umstände gewissermaßen “gezwungen” wird, die Assoziation von negativem Affekt mit dem Anreizobjekt 000 zu speichern, scheint eine Übertreibung der durch den Hippocampus vorgenommenen Minimierung von Überlappungen in den Assoziationen zu sein. Dieses Problem, das nicht ohne weiter gehende theoretische Überlegungen untersucht werden kann, soll im folgenden Abschnitt behandelt werden.

#### 4.4 Theoretische Überlegungen zu den Ergebnissen des Simulationsexperiments

In diesem Abschnitt soll noch näher auf die Inhalte des Extensionsgedächtnisses in den verschiedenen Gruppen eingegangen werden. Im wesentlichen ergeben sich zwei Deutungsmöglichkeiten, die auf den ersten Blick widersprüchlich erscheinen mögen, sich jedoch in Wirklichkeit ergänzen. Hierbei wird davon ausgegangen, dass sich sämtliche Ergebnisse, wie in Abschnitt 4.3 erläutert, allein auf den Mechanismus der Informationsübertragung ins Extensionsgedächtnis zurückführen lassen. Dass teilweise Assoziationen übertragen werden, die nicht plausibel erscheinen — hier sei vor allem auf die in einem eigenen Unterabschnitt behandelte Kombination von  $A--(000)$  und  $A+-Ann-(100)$  hingewiesen — lässt einerseits die Deutung zu, dass die Integration neuer Inhalte ins Extensionsgedächtnis noch durch andere kognitive Systeme unterstützt werden sollte. Hier wird eine Unterstützung des *Fühlens* durch das *Denken* vorgeschlagen. Andererseits können sich auch auf den ersten Blick unplausible Assoziationen als adaptiv erweisen — und mit dem Begriff der *Selbstmotivierung* in Zusammenhang gebracht werden.

##### Selbstmotivierung

Betrachtet man die paradoxe Kombination der Assoziationen  $A--(000)$  und  $A+-Ann-(100)$ , so liegt es nahe, dies einfach als eine “Fehleinschätzung” anzusehen; oder besser: als Schwäche im Mechanismus der Informationsübertragung ins Extensionsgedächtnis. Eine andere Möglichkeit, die Assoziation  $A+-Ann-(100)$  zu interpretieren, führt zu einer interessanten Perspektive



und

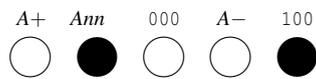
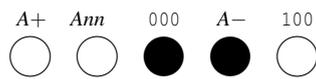


Abbildung 4.19: Die bestrafungssensitiven Persönlichkeitstypen übertragen auch Verknüpfungen mit negativem Affekt ins Extensionsgedächtnis. Dadurch haben sie zwei Möglichkeiten, die Assoziation Ann-(100) zu kombinieren: Einerseits mit A+-(000), andererseits jedoch mit A--(000) — beide Möglichkeiten haben einen Overlap (2.15) von -3.

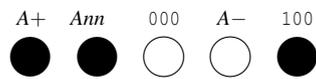
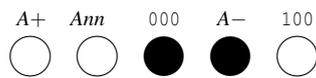


Abbildung 4.20: Bei Ahnungsvollen kommt die häufigste Kombination von Assoziationen durch ein Zusammenwirken der Affektsensibilitäten und der Minimierung des Overlaps zustande: Wegen der Sensibilität für positiven Affekt kann 100 nur in Kombination mit positivem Affekt ins Extensionsgedächtnis übertragen werden: Positiver Affekt, der zum Ausführen des Annäherungsverhaltens benötigt wird, kann bis zur Übertragung ins Extensionsgedächtnis nicht herausgedämpft werden. Andererseits führt die Sensibilität für negativen Affekt dazu, dass häufig 000 mit negativem Affekt verknüpft werden kann. Wegen des minimalen Overlaps (2.15) von -5 werden diese beiden Assoziationen, sobald sie auftreten, gemeinsam im Extensionsgedächtnis gespeichert.

---

auf die *Selbstmotivierung* (siehe Abschnitt 1.2): Oben wurde erklärt (siehe Abschnitt 4.3), wie die im Zusammenhang mit der Objektkategorie 100 sicher nicht die “Realität” repräsentierende Assoziation mit Annäherungsverhalten *und* positivem Affekt zustande kommt. Zu beachten ist aber, dass dies in der Situation unseres Experiments die besten Voraussetzungen dafür schafft, das Objekt der Kategorie 111 zu erreichen: Die Prämisse `GetPosAff` wird genau dann aktiv, wenn eine Abfrage des Extensionsgedächtnisses mit *Ann*-(100) positiven Affekt zurückliefert. Das Extensionsgedächtnis hat also für das “unangenehme” Ziel 111 eine Assoziation zu positivem Affekt hergestellt. Es wird damit positiver Affekt mit genau demjenigen Objekt assoziiert, dessen Ansteuern möglicherweise zusätzliche Motivation erfordert. Wird die Objektkategorie 111 — wie zu Beginn des Abschnitts — als “unangenehmes”, schwierig zu erreichendes Ziel interpretiert, jedoch nicht als tatsächlich “gefährlich” für die Person, erscheint die vorgenommene *Umbewertung* durchaus sinnvoll. Eine solche Interpretation des Objektes 111 scheint durchaus gerechtfertigt, da in der Konzeption des Experiments keine Möglichkeit vorgesehen war, das Zielobjekt zu wechseln. In einer realen Situation würden im Falle eines tatsächlich als bedrohlich wahrgenommenen Objekts, die elementaren Systeme der Objekterkennung und Verhaltenssteuerung Fluchtverhalten in Gang setzen. Das Speichern der Assoziation *A+Ann*-(100) erinnert also an die *kongruenzorientierte Aufmerksamkeit* für Erfolgserlebnisse beim Erreichen eines schwierigen Ziels: Beim Ansteuern des “unangenehmen” Objekts 111 wird bei der Übertragung ins Extensionsgedächtnis nur positiver Affekt berücksichtigt. Dies führt dazu, dass die Motivation zum Erreichen des Ziels gesteigert wird. Wie in Abschnitt 4.3 dargestellt, ergibt sich dieser Effekt einzig und allein dadurch, dass das Extensionsgedächtnis nur in entspannter Affektlage zur Informationsaufnahme bereit ist. Es stellt sich also die Frage, ob *Selbstmotivierung* sich zumindest teilweise *automatisch* aus dem in Abschnitt 1.5 erläuterten Mechanismus ergibt. In diesem Fall wäre *Selbstmotivierung* eine inhärente Eigenschaft des Extensionsgedächtnisses, die ohne zusätzliche Mechanismen oder gar eine bewusste, intentionale Steuerung zustande kommt.

### **Leistungsdefizite der bestrafungssensitiven und ahnungsvollen Typen**

Bei den *bestrafungssensitiven* Typen sowie die *ahnungsvolle* wird die als möglicher Kern der *Selbstmotivierung* ausgemachte *Umbewertung* der Objektkategorie 100 übertrieben: Paradoxerweise wird hier eher die bedürfnisbefriedigende Objektkategorie 000 mit negativem Affekt besetzt. Damit ergibt sich ein Funktionsdefizit des Selbst, das möglicherweise mit der Situation des von Kuhl beschriebenen depressiven Patienten in Verbindung gebracht werden, der von seiner Therapeutin vergeblich aufgefordert wird, “an seine Möglichkeiten zur Selbstbestimmung zu glauben. Diese Therapeutin handelt paradox, wenn die primäre Ursache der Depression in diesem konkreten Fall nicht in einem falschen Wollen oder in ungünstigen Denkinhalten, sondern in einem funktionalen Defizit begründet ist: Wenn der Patient nicht wollen und nicht glauben *kann*, wenn also ein Funktionsdefizit vorliegt, wird die implizite Verantwortungszuschreibung u.U. nicht die erwartete aktivierende Wirkung haben, sondern die Hoffnungslosigkeit noch weiter vergrößern.” [9, Kuhl, S.43] Im Gegensatz zu Kuhls Interpretation liegt im vorliegenden Experiment das “Defizit” der *bestrafungssensitiven* bzw. *ahnungsvollen* Persönlichkeitstypen nicht in einem fehlenden Zugang zu Selbstrepräsentationen, sondern in einer sich zwangsläufig ergebenden Fehlerhaftigkeit der Repräsentationen im Extensionsgedächtnis. Es wird deshalb die abschließende Frage aufgeworfen, ob der bisher rein empirische Prozess der Integration von Wissen ins Extensionsgedächtnis schon vollständig ist oder noch durch eine analytische Überprüfung der ins Extensionsgedächtnis übertragenden Inhalte ergänzt werden muss.

---

## Übersteigter Empirismus?

Das Zustandekommen der Kombination von Assoziationen wie  $A--(000)$  mit  $A+-Ann-(100)$  kann als Schwäche der Informationsübertragung ins Extensionsgedächtnis gesehen werden — immerhin wird der Zugang zu einem bedürfnisbefriedigenden Objekt versperrt —, die sich aus einer noch nicht vollständig modellierten Informationsübertragung ins Extensionsgedächtnis ergibt. An dieser Stelle zeigt sich offensichtlich eine Schattenseite unseres “Humeschen Wesens”; die so einfache Bildung von Assoziationen allein unter dem Kriterium *Berührung in Raum und Zeit* stößt an ihre Grenzen, das Sammeln von Erfahrungen zur einzigen Grundlage des *Fühlens* zu machen, erweist sich als möglicherweise nicht ausreichend. Immanuel Kant reagierte auf David Humes radikalen Empirismus, indem er für den Erkenntnisgewinn einen Mittelweg zwischen reinem *Empirismus* und reinem *Rationalismus* vorschlug (Rationalismus wird hier als Betonung der Vernunft als Quelle des Erkenntnisgewinns gegenüber der Sinneserfahrung verstanden<sup>20</sup>). Für uns stellt sich also die Frage: Wie lässt sich — übertragen auf die PSI-Theorie — der *Empirismus* des Extensionsgedächtnisses durch eine *rationale* Komponente ergänzen?

## Unterstützung des *Fühlens* durch das *Denken*

*Rational* im Sinne der obigen Definition arbeitet das kognitive System des *analytischen Denkens*. Deshalb sollte — bei einer Erweiterung des vorliegenden Modells um das Denken — nicht nur untersucht werden, wie das Fühlen das Denken beeinflusst<sup>21</sup>, sondern auch, in welchem Maß das Denken Einfluss auf das Sammeln von Erfahrungen durch das Fühlen nimmt. Hier sollten vor allem die Fälle betrachtet werden, in denen sich ein Widerspruch zu bereits bestehendem Erfahrungswissen ergibt — das analytische Denken ist das prädestinierte System für das Entdecken von Widersprüchen.

## Integration von Gegensätzen

Laut der PSI-Theorie sollten im Extensionsgedächtnis Widersprüche allerdings nicht “korrigiert”, sondern integriert werden<sup>22</sup>: Eine durch das Denken gefundene, zu den Inhalten des Extensionsgedächtnisses widersprüchliche Assoziation, wird normalerweise die bestehende Assoziation nicht “überschreiben”, stattdessen sollten beide Assoziationen im Extensionsgedächtnis gespeichert werden. Dies steht in Konflikt mit der Minimierung des Ähnlichkeitsmaßes (2.19): Widersprüchliche Assoziationen stimmen wenigstens in dem Merkmal überein, zu dem sie widersprüchlich Verknüpfungen bilden. Zusätzlich zu der Übertragung möglichst wenig übereinstimmender Assoziationen von der Objekterkennung, sollte eine Übertragungsmöglichkeit vom Denkens ins Extensionsgedächtnis beim Auftreten von Widersprüchen vorgesehen werden.

---

<sup>20</sup>Johannes Hirschberger schreibt in seiner “Geschichte der Philosophie” über den *Rationalismus* in Gegensatz zum *Empirismus*: “Ein prinzipieller Unterschied besteht in der Bewertung der Erfahrung. Gemeint ist immer nur die Sinneserfahrung. Für den Rationalismus ist sie nur Material, Gelegenheit, Einleitung. Wissenschaft und Wahrheit eröffnet sich im Geist und seinen notwendigen Wesenseinsichten.” [4, Hirschberger, S.188]

<sup>21</sup>Zahlreiche Einflüsse des Fühlens auf das Denken sind in [9, Kuhl, Kapitel 13] dargestellt.

<sup>22</sup>Siehe Fußnote 11, 26

# Kapitel 4

## Fazit

Mit dem vorliegenden Modell gelingt es, wichtige Aussagen der PSI-Theorie zur Verarbeitungsform des Fühlens durch ein einfaches Modell zu reproduzieren — es ergaben sich abhängig von Affektsensibilitäten unterschiedliche Inhalte des Extensionsgedächtnisses; dabei wies das Extensionsgedächtnis den postulierten positiven Bias auf und auch die Leistungsunterschiede der verschiedenen Parametrisierungen entsprachen den Erwartungen.

Das besondere Potential der Durchführung von Simulationsexperimenten zeigte sich schon im hier durchgeführten, einfachen Experiment: Es ergab sich ein interessanter Erklärungsansatz für die Entstehung der *Selbstmotivierung* und die Leistungsdefizite von *bestrafungssensitiven* und *ahnungsvollen* Persönlichkeitstypen.

Allerdings muss sich der Autor hier die Frage gefallen lassen, ob das Experiment in den sich anschließenden theoretischen Überlegungen nicht überinterpretiert wurde. Dem ist entgegenzuhalten, dass sich — wie in der Auswertung gezeigt wurde — bestimmte für die Person ungünstige Inhalte des Extensionsgedächtnisses *zwangsläufig* durch den postulierten Prozess der Informationsübertragung ergeben. Dies zeigt, dass das vorliegende Modell noch nicht alle Einzelheiten des Fühlens abbildet: Kombinationen von Assoziationen wie  $A-(000)$  und  $A+-Ann-(100)$ , die der Person den Zugang zu bedürfnisbefriedigenden Objekten verwehren, ergeben sich mehr oder weniger automatisch durch die Minimierung des Ähnlichkeitsmaßes (2.19). Keine Berücksichtigung findet dagegen bisher die Integration von Inhalten, die mit bereits bekanntem Erfahrungswissen im Widerspruch stehen. Es wurde vorgeschlagen, dem noch zu modellierenden kognitiven System des *Denkens* diese Integration von Widersprüchen ins Extensionsgedächtnis zu übertragen.

Unabhängig von der Richtigkeit der aus dem Experiment entwickelten Ideen, zeigen sie doch zumindest, dass das vorliegende Modell interessante Anstöße für die nähere Betrachtung von Konzepten der PSI-Theorie zu geben vermag.

# Anhang A

## Resultate des Simulationsexperiments

### 1. Einzelne im EG gespeicherte Assoziationen

**Tabelle A.1: Einzelne Assoziationen in 27 Simulationsläufen  
Gruppe 1: Belohnungssensitive Typen**

Persönlichkeitsstil	A+- <i>(000)</i>	A+- <i>Ann</i> - <i>(000)</i>	A+	A+- <i>Ann</i> - <i>(100)</i>	$\varnothing t_{end} (n_{end})$
liebenswert	20	4	6	26	131.2(27)
selbstbestimmt	20	4	6	26	131.2(27)
ehrgeizig	22	5	5	26	84.3(27)

**Tabelle A.2: Einzelne Assoziationen in 27 Simulationsläufen  
Gruppe 2: Bestrafungssensitive Typen**

Persönlichkeitsstil	A+- <i>(000)</i>	A+- <i>Ann</i> - <i>(100)</i>	<i>Ann</i> - <i>(000)</i>	<i>Ann</i> - <i>(100)</i>	A-- <i>(000)</i>	A-- <i>Ann</i> - <i>(111)</i>	nichts	$\varnothing t_{end} (n_{end})$
sorgfältig	14(12)	2(1)	1(1)	22(12)	14(3)	2(0)	1(1)	130.4(14)
selbstkritisch	12(11)	1(1)	1(1)	24(14)	16(6)	1(0)	1(1)	147(16)
zurückhaltend	10(7)	3(2)	1(1)	20(9)	18(6)	3(0)	1(1)	128.4(12)

**Tabelle A.3: Einzelne Assoziationen in 27 Simulationsläufen  
Gruppe 3: Eigenwillige Persönlichkeit**

Persönlichkeitsstil	A+- <i>(000)</i>	A+- <i>Ann</i> - <i>(000)</i>	<i>Ann</i> - <i>(000)</i>	<i>Ann</i> - <i>(100)</i>	nichts	$\varnothing t_{end} (n_{end})$
eigenwillig	25	1	1	25	3	139.6(27)

### 2. Kombinationen der im EG gespeicherten Assoziationen

**Tabelle A.4: Einzelne Assoziationen in 27 Simulationsläufen  
Gruppe 4: Ahnungsvolle Persönlichkeit**

Persönlichkeitsstil	A+- (000)	A+-Ann- (100)	A+	A-- (0)	A-	nichts	$\otimes t_{end} (n_{end})$
ahnungsvoll	3(3)	24 (10)	3(0)	21(4)	2(2)	1(1)	63.7(10)

**Tabelle A.5: Stabile Kombinationen in 27 Simulationsläufen  
Gruppe 1: Belohnungssensitive Persönlichkeitstypen**

Persönlichkeitsstil	A+ A+-Ann- (000)	A+ A+-Ann- (100)	A+ A+-Ann- (000) A+-Ann- (100)	A+- (000) A+-Ann- (100)	A+-Ann- (100) (4x)	$\otimes t_{end} (n_{end})$
liebenswert	1	2	3	20	1	131.2 (27)
selbstbestimmt	1	2	3	20	1	131.2(27)
ehrgeizig	1		4	22		84.3(27)

**Tabelle A.6: Stabile Kombinationen in 27 Simulationsläufen  
Gruppe 2: Bestrafungssensitive Typen**

Persönlichkeitsstil	A+- (0) Ann- (100)	A-- (0) A+-Ann- (100)	A-- (0) Ann- (111)	nichts A+- (000) A-- (000) Ann- (000)	$\otimes t_{end} (n_{end})$
sorgfältig	11	1	1	1	130.4(14)
selbstkritisch	10	1	4	1	147(16)
zurückhaltend	6	2	3	1	128.4(12)

Persönlichkeitsstil	A+- (000) A--Ann- (111)	A-- (000) Ann- (100)	A-- (000) A+-Ann- (100)
sorgfältig	2	10	1
selbstkritisch	1	10	
zurückhaltend	3	11	1

**Tabelle A.7: Stabile Kombinationen in 27 Simulationsläufen  
Gruppe 3: Eigenwillige Persönlichkeit**

Persönlichkeitsstil	A+- (000) Ann- (100)	nichts A+-Ann- (000)	nichts Ann- (100)	nichts A+- (000) Ann- (000)	$\otimes t_{end} (n_{end})$
eigenwillig	24	1	1	1	139.6

**Tabelle A.8: Stabile Kombinationen in 27 Simulationsläufen  
Gruppe 4: Ahnungsvolle Persönlichkeit**

Persönlichkeitsstil	$A+-(000)$	$A-$	$A--(000)$	nichts	$\otimes t_{end} (n_{end})$
	$A+-Ann-(100)$	$A+-Ann-(100)$	$A+-Ann-(100)$	$A+-Ann-(100)$	
ahnungsvoll	3	2	4	1	63.7(10)
	$A+$	$A--(000)$			
	$A--(000)$	$A+-Ann-(100)$			
ahnungsvoll	3	14			

# Literaturverzeichnis

- [1] Baehni, Sebastien (2000) *Neural JAVA*, Tutorial des “Laboratoire de Calcul Neuro-mimétique (LCN)” der “École Polytechnique Fédérale de Lausanne”, Lausanne  
XOR-Problem: <http://diwww.epfl.ch/w3mantra/tutorial/english/perceptron/html/index.html>  
Hopfield-Netz: <http://diwww.epfl.ch/w3mantra/tutorial/english/hopfield/html/index.html>,  
26.10.2004
- [2] Bröcker, Thomas (2000) *Adipositas, Metabolisches Syndrom, Neuropathie, Herzfrequenzvariabilität, Emotionales Befinden, Persönlichkeitsdispositionen und ein PSI-Modell*, Dissertation, Universität Osnabrück, S.53-77  
<http://elib.ub.uni-osnabrueck.de/elib/user/>,  
oder in [9, S.1032-1063]
- [3] Hammer, Barbara (2000) *Neuronale Netze*, Vorlesungsskript für das Wintersemester 1999/2000
- [4] Hirschberger, Johannes (1980) *Geschichte der Philosophie*, hier zitiert nach: Nachdruck der elften Auflage (Herder Verlag, Freiburg im Breisgau), Lizenzausgabe für KOMET, Frechen
- [5] Hume, David (1967) *Eine Untersuchung über den menschlichen Verstand*, Übersetzung des Originalwerks *An Enquiry Concerning Human Understanding*, Reclam, Stuttgart
- [6] Jätzold, Stephan *Quicksilver-Homepage* <http://www.usf.uos.de/projects/quicksilver/>
- [7] Jefferson Airplane, *White Rabbit*, <http://www.mp3lyrics.org>, 1.Juni 2004
- [8] Jung, C. G. (1936/1990) *Typologie*, dtv
- [9] Kuhl, Julius (2001) *Motivation und Persönlichkeit - Interaktionen psychischer Systeme*, Hogrefe, Göttingen
- [10] McClelland, J. L., McNaughton, B. L. & O'Reilly, R. C. (1995) *Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory*, *Psychological Review*, 102, 419–457
- [11] David E. Rumelhart, James L. McClelland & the PDP Research Group (1986) *Parallel distributed processing : explorations in the microstructure of cognition*, Volume 1: Foundations, The MIT Press

- 
- [12] Schacter, D.L. (2004), aus: *Project on the decade of the brain*, Zitat des Artikels *Memory: The Fragile Power*, Library of Congress,  
<http://www.loc.gov/loc/brain/emotion/Schacter.html>, 26.10.2004
- [13] in Malchow, Horst (Hrsg.) (2001) *Modellbildung und -anwendung in den Wissenschaften V*, Beiträge des Instituts für Umweltsystemforschung (Hrsg.: Matthies, Michael), Nr.23, Universität Osnabrück, S.7-30 <http://www.usf.uos.de/usf/beitraege/>,  
26.10.2004
- [14] Siekmann, Ivo (2003) *Alice im Wunderland – Ein agentenbasiertes PSI-Modell*, Projektbericht zum Wochenprojekt im Anschluss an die Vorlesung *Agentenbasierte Modellierung*, <http://www.usf.uos.de/projects/quicksilver/model-archive/abm-2003/psi.pdf>,  
26.10.2004
- [15] Siekmann, Ivo (2003) *Alice im Wunderland – Ein agentenbasiertes PSI-Modell*, Projektbericht zum zweimonatigen Studienprojekt, unveröffentlicht.
- [16] Tarantino, Quentin (1994) *Pulp Fiction*,  
<http://www.godamongdirectors.com/scripts/pulp.shtml>, 6.Juni 2004
- [17] Tolman, E.C. (1948) *Cognitive maps in rats and men*, *Psychological Review*, 55, 189-208
- [18] Traeger, Dirk H. (1993) *Einführung in die Fuzzy-Logik*, Teubner, Stuttgart
- [19] Volgushev, Maxim (2004) *Hauptvorlesung Physiologie WS 2003/2004*, RUHR-UNIVERSITÄT BOCHUM, INSTITUT FÜR PHYSIOLOGIE, Medizinische Fakultät,  
<http://www.neurop.ruhr-uni-bochum.de/~maxim/DownLoads/LernenGedMax04.pdf>
- [20] Paul Watta, Mohamad Hassoun & Norman Dannug (1996) *A Backprop Learning Tool for Function Approximation*,  
<http://neuron.eng.wayne.edu/bpFunctionApprox/bpFunctionApprox.html>,  
26.10.2004
- [21] Zell, Andreas (1994) *Simulation Neuronaler Netze*, 1.Auflage, Addison-Wesley, Stuttgart
-

---

---

---

## **Beiträge des Instituts für Umweltsystemforschung der Universität Osnabrück**

1. Eberhard Umbach: Umweltverträgliches Wirtschaftssystem in den Bereichen Abfall und Emissionen. März 1997.
  2. Stefan Trapp, Bernhard Reiter, Michael Matthies: Überprüfung und Fortentwicklung der Bodenwerte für den Boden-Pflanze-Pfad - Teilprojekt Transferfaktoren Boden-Pflanze. August 1997.
  3. Michael Matthies (Hrsg.): Stoffstromanalyse und Bewertung. September 1997.
  4. Dirk Melcher: Quantifizierung, Klassifizierung und Modellierung der Phytotoxizität organischer Chemikalien. Oktober 1997.
  5. Stefan Schwartz: Organische Schadstoffe in der Nahrungskette - Vorstudie zur Validierung von Expositionsmodellen. November 1997.
  6. Volker Berding: Private Hausbrunnen - Vergleichende Bewertung von Maßnahmen zur Verbesserung der Trinkwasserqualität. Oktober 1997.
  7. Horst Malchow (Hrsg.): Modellbildung und -anwendung in den Wissenschaften I. Januar 1998.
  8. Birgit Radtke: Bifurkationen in einem Modell mariner Planktodynamik. Januar 1998.
  9. Werner Berens: Konzeption eines Umweltinformationssystems für die Universität Osnabrück. Juni 1998.
  10. Michael Matthies (Hrsg.): Studienprojekte 1998. September 1998.
  11. Michael Matthies (Hrsg.): Globaler Wandel. September 1998.
  12. Klaus Brauer (Hrsg.): Institutsbericht. September 1998.
  13. Klaus Brauer, Horst Malchow, Michael Matthies, Eberhard Umbach (Hrsg.): Materialien des Arbeitstreffens Systemwissenschaft in der Lehre, Universität Osnabrück, 29./30.9.1998. Dezember 1998.
  14. Horst Malchow (Hrsg.): Modellbildung und -anwendung in den Wissenschaften II. Dezember 1998.
  15. Horst Malchow (Hrsg.): Modellbildung und -anwendung in den Wissenschaften III. August 1999.
  16. Michael Matthies (Hrsg.): Regionale Nachhaltigkeit. September 2000.
  17. Markus Klein: Langjähriger Wasserhaushalt von Gras- und Waldbeständen. Entwicklung, Kalibrierung und Anwendung des Modells LYFE am Groß-Lysimeter St. Arnold. Juni 2000.
-

- 
18. Markus Brune: Multimediale Umweltmodellierung mit Fuzzy-Mengen. Juli 2000.
  19. Michael Matthies (Hrsg.): Fraktale in Hydrologie und Biologie. Oktober 2000.
  20. Stefan Fuest (Dissertation): Regionale Grundwassergefährdung durch Nitrat. Dezember 2000.
  21. Carsten Schulze (Dissertation): Modelling and evaluating the aquatic fate of detergents. Januar 2001.

Die Beiträge können gegen einen Selbstkostenpreis (ca. 10 € pro Exemplar) beim Institut für Umweltsystemforschung, Universität Osnabrück, 49069 Osnabrück bestellt werden.

Alle folgenden Beiträge sind herunterzuladen unter <http://www.usf.uos.de/usf/beitraege/>.

22. Horst Malchow (Hrsg.): Modellbildung und -anwendung in den Wissenschaften IV. Januar 2001.
  23. Horst Malchow (Hrsg.): Modellbildung und -anwendung in den Wissenschaften V. August 2001.
  24. Kai Leßmann (Diplomarbeit): Probabilistic Exposure Assessment. Parameter Uncertainties and their Effects on Model Output. November 2002.
  25. Frank M. Hilker (Diplomarbeit): Parametrisierung von Metapopulationsmodellen. März 2003.
  26. Nadja Rüger (Diplomarbeit): Habitat suitability for *Populus euphratica* in the Northern Amudarya delta - a fuzzy approach. Juni 2003.
  27. Claudia Pahl-Wostl, Eva Ebenhöf (Hrsg.): Komplexe Adaptive Systeme. Juli 2003.
  28. Horst Malchow (Hrsg.): Chaos und Ordnung in Natur und Gesellschaft. Dezember 2004.
  29. Andreas Focks (Diplomarbeit): Modeling the transfer of antibiotic drug resistance genes between *E. coli* strains. Juni 2005.
  30. Christiane Zarfl (Diplomarbeit): Modellierung von Arsen in der Mulde. Juni 2005.
  31. Sven Lautenbach (Dissertation): Modellintegration zur Entscheidungsunterstützung für die Gewässergütebewirtschaftung im Einzugsgebiet der Elbe. November 2005.
  32. Frank M. Hilker and Frank H. Westerhoff: Control of chaotic population dynamics: Ecological and economic considerations. November 2005.
  33. Harold Fellermann (Diplomarbeit): Micelles as containers for protocells. Dezember 2005.
  34. Jens Newig, Oliver Fritsch (Hrsg.): Effektivität von Entscheidungsprozessen. Mai 2006.
  35. Ba Kien Tran (Diplomarbeit): Modellierung biologischer Invasionen mit Reaktions-Diffusionsgleichungen. Juli 2006.
  36. Ivo Siekmann (Diplomarbeit): Agentenbasierte Modellierung von Persönlichkeitsunterschieden auf der Grundlage der PSI-Theorie. Juli 2006.
-

