

# Informatik im Kontext 1

Grundlagen der Mensch-Computer-Interaktion

Frank Steinicke      Kim Wittenburg

Version 1.0



# Inhaltsverzeichnis

<b>1</b>	<b>Motivation</b>	<b>6</b>
1.1	Das PACT-Framework . . . . .	6
1.2	Ursprünge der MCI . . . . .	7
1.3	Warum HCI? . . . . .	9
<b>2</b>	<b>Menschliche Informationsverarbeitung</b>	<b>11</b>
2.1	Benutzerschnittstellen . . . . .	11
2.2	Informationspsychologie . . . . .	12
2.3	Modelle der menschlichen Informationsverarbeitung . . . . .	16
2.3.1	Model Human Processor . . . . .	16
2.3.2	Soar . . . . .	17
2.3.3	Weitere Modelle . . . . .	18
2.4	Prättentive und attentive Merkmale . . . . .	18
<b>3</b>	<b>Wahrnehmung</b>	<b>21</b>
3.1	Wahrnehmungstheorie . . . . .	21
3.2	Physikalische Grundlagen . . . . .	24
3.3	Visuelle Wahrnehmung . . . . .	25
3.3.1	Physikalische Hintergründe . . . . .	25
3.3.2	Biologische Hintergründe . . . . .	28
3.3.3	Sehschärfe . . . . .	32
3.3.4	Größenwahrnehmung . . . . .	32
3.3.5	Farbmodelle . . . . .	33
3.3.6	Helligkeits- und Kontrastwahrnehmung . . . . .	35
3.3.7	Objekterkennung . . . . .	37
3.3.8	Gesichtserkennung . . . . .	40
3.4	Auditive Wahrnehmung . . . . .	41
3.4.1	Physikalische Hintergründe . . . . .	41
3.4.2	Biologische Hintergründe . . . . .	42
3.4.3	Geräuschlokalisierung . . . . .	43
3.4.4	Akustik in der MCI . . . . .	43
3.4.5	Vestibulärer Sinn . . . . .	44
3.5	Tastsinn . . . . .	45
3.5.1	Propriozeption . . . . .	45
3.6	Geschmacks- und Geruchssinn . . . . .	46
3.7	Psychophysik . . . . .	46
3.7.1	Weber-Fechner-Gesetz . . . . .	47

3.7.2	Adaption . . . . .	47
3.7.3	Habituation . . . . .	47
3.7.4	Sensorische Integration . . . . .	48
<b>4</b>	<b>Kognition</b>	<b>50</b>
4.1	Gedächtnis . . . . .	50
4.1.1	Sensorische Register . . . . .	51
4.1.2	Arbeitsgedächtnis . . . . .	51
4.1.3	Langzeitgedächtnis . . . . .	52
4.2	Lernen . . . . .	54
4.2.1	Lernmethoden . . . . .	55
4.2.2	Fertigkeitserwerb . . . . .	55
4.2.3	Lernform . . . . .	56
4.2.4	Lernbewertung . . . . .	56
4.2.5	Potenzgesetz der Übung . . . . .	57
4.3	Logisches Denken . . . . .	57
4.3.1	Deduktion . . . . .	58
4.3.2	Induktion . . . . .	59
4.3.3	Abduktion . . . . .	59
4.4	Entscheidungen treffen (Gesetz von Hick) . . . . .	60
4.5	Aufmerksamkeit und Belastung . . . . .	61
4.5.1	Aufmerksamkeit . . . . .	61
4.5.2	Belastung . . . . .	62
<b>5</b>	<b>Kommunikation und Handlung</b>	<b>64</b>
5.1	Kommunikation . . . . .	64
5.1.1	Kommunikationssicht . . . . .	65
5.1.2	Interaktionssicht . . . . .	66
5.2	Handlung . . . . .	66
5.2.1	Human Action Cycle . . . . .	68
5.3	Fehler . . . . .	68
5.3.1	Fehlerklassifikation . . . . .	70
5.4	Aktion und Motorik . . . . .	71
5.4.1	Biologischer Hintergrund . . . . .	71
5.4.2	Reafferenzprinzip . . . . .	72
5.5	Modelle . . . . .	72
5.5.1	GOMS . . . . .	73
5.5.2	Fitts' Law . . . . .	75
5.5.3	Steering Law . . . . .	77
5.5.4	CLC-Model . . . . .	78
5.6	Kontrolle-Display-Relation . . . . .	78
5.6.1	Räumliche Relation . . . . .	79
5.6.2	C/D-Sensitivität . . . . .	80
5.6.3	Three-State-Model . . . . .	81

<b>6</b>	<b>Historie der MCI</b>	<b>82</b>
6.1	Generationen der Informatik . . . . .	82
6.1.1	Hardware . . . . .	82
6.1.2	Software . . . . .	86
6.1.3	Mensch-Computer-Interaktion . . . . .	87
6.2	Klassische Eingabegeräte . . . . .	88
6.2.1	Maus . . . . .	88
6.2.2	Tastatur . . . . .	90
<b>7</b>	<b>Interaktionsparadigmen</b>	<b>93</b>
7.1	Klassische Interaktionsgestaltung . . . . .	93
7.1.1	Kommandosysteme . . . . .	93
7.1.2	Form-Fill Interfaces . . . . .	94
7.1.3	Point & Click Interfaces . . . . .	94
7.1.4	Menu Interfaces . . . . .	95
7.1.5	Direkte Manipulation . . . . .	95
7.2	WIMP-GUIs . . . . .	96
7.2.1	Windows . . . . .	96
7.2.2	Icons . . . . .	97
7.2.3	Menus . . . . .	97
7.2.4	Pointer . . . . .	98
7.2.5	Widgets . . . . .	98
7.2.6	Look-&-Feel . . . . .	98
7.3	Weitere Interface-Type . . . . .	98
<b>8</b>	<b>Interaktionsdesign</b>	<b>102</b>
8.1	Leitsatz für Gutes Design . . . . .	102
8.2	Affordanzen . . . . .	103
8.2.1	Reale Affordanzen . . . . .	103
8.2.2	Virtuelle Affordanzen . . . . .	105
8.3	Mentale & Konzeptionelle Modelle . . . . .	105
8.3.1	Mentale Modelle . . . . .	105
8.3.2	Konzeptionelle Modelle . . . . .	106
8.3.3	Mappings . . . . .	107
8.4	Design-Regelwerke . . . . .	108
8.4.1	Prinzipien . . . . .	109
8.4.2	Heuristiken/Goldene Regeln . . . . .	111
8.4.3	Patterns . . . . .	112
8.4.4	Standards . . . . .	113
8.4.5	Guidelines . . . . .	113
<b>9</b>	<b>Evaluierung</b>	<b>115</b>
9.1	Arten der Evaluierung . . . . .	115
9.1.1	Warum? . . . . .	115

9.1.2	Wann? . . . . .	115
9.1.3	Feld- und Laborstudien . . . . .	116
9.1.4	Ergebnisse . . . . .	117
9.2	Analytische Methoden . . . . .	118
9.3	Empirische Methoden . . . . .	118
9.4	Datenanalyse . . . . .	119
9.4.1	Deskriptive Statistik . . . . .	120
9.4.2	Subgruppenanalyse . . . . .	120
9.4.3	Grafiken . . . . .	121
9.5	Statistische Tests . . . . .	121
9.6	DECIDE-Framework . . . . .	122

<b>Lösungen zu den Aufgaben</b>	<b>124</b>
---------------------------------	------------

# 1 Motivation

Dies ist das Skript zur Veranstaltung *IKON-1: Grundlagen der Mensch-Computer-Interaktion*, einer der beiden Teilveranstaltungen des Moduls *IKON: Informatik im Kontext*. In dieser Veranstaltung geht es darum einen allgemeinen Überblick über das Forschungsgebiet der Mensch-Computer-Interaktion zu gewinnen. Dazu werden wir sowohl den Menschen als auch den Computer eingehend analysieren und verschiedene Interaktionsparadigmen kennen lernen.

Es ist sehr wichtig zu verstehen, dass die Mensch-Computer-Interaktion eine stark interdisziplinäre Wissenschaft ist. Es reicht nicht, ein Problem nur aus der Sicht der Informatik zu betrachten. Meistens muss man ebenso Ergebnisse aus der Psychologie, den Gestaltungswissenschaften, den Kommunikationswissenschaften oder anderen Wissenschaften anwenden. Im Laufe der Veranstaltung werden wir verschiedene Disziplinen kennen lernen und anwenden, um den Menschen und die Mensch-Computer-Interaktion besser zu verstehen.

## 1.1 Das PACT-Framework

PACT-Framework

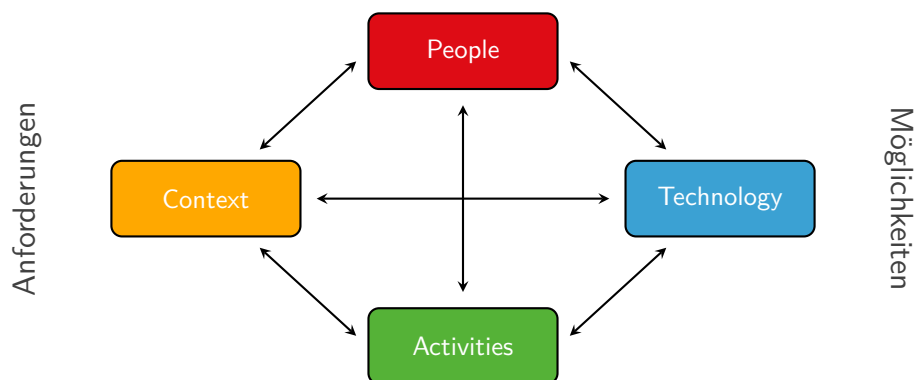


Abbildung 1.1: PACT-Framework

Das PACT-Framework (Abbildung 1.1) ist eines der grundlegenden Modelle der *Informatik im Kontext*. Es besteht aus vier Komponenten: Menschen (*engl. People*), Aufgaben (*engl. Activities*), Kontext (*engl. Context*) und Technologie (*engl. Technology*). In einem Satz lässt sich das PACT-Framework folgendermaßen charakterisieren: Heterogene **Menschen** führen verschiedenste **Aufgaben und Aktivitäten** in diversen **Kontexten** mit Hilfe von unterschiedlichen **Technologien** durch. Dabei wechselwirken die vier Komponenten ständig miteinander. Beispielsweise beeinflusst die Einführung

neuer, interaktiver Systeme Prozesse in der realen Welt. Besonders sieht man das zum Beispiel bei Smartphones, die innerhalb von wenigen Jahren zu einem allgegenwärtigen Teil der Gesellschaft geworden sind. Im Sinne von PACT hat hier also eine Wechselwirkung aus der Technologie Smartphone im Kontext Gesellschaft stattgefunden, die sich auf Menschen und deren Aktivitäten ausgewirkt hat, bzw. auswirkt. Natürlich kann die Einführung neuer Technologien ebenso negative Auswirkungen haben, die sich ebenfalls mittels des PACT-Frameworks darstellen lassen. Außerdem sind neue Technologien auch oft mit Lernkurven oder der Rationalisierung von Arbeitsabläufen verbunden.

Anwendung findet das PACT-Framework beim Design und der Entwicklung neuer Technologien. Damit auch wirklich ein Produkt entwickelt wird, das den Anforderungen der Kunden entspricht, müssen die Kunden selbst, aber eben auch deren Aktivitäten und der Einsatzraum des Produktes berücksichtigt werden. Daher sollte man immer, wenn man an einer neuen Technologie arbeitet, das PACT-Framework bedenken.

### Aufgabe 1

In Abbildung 1.1 zum PACT-Framework ist die Seite *Context* mit „Anforderungen“ und die Seite *Technology* mit „Möglichkeiten“ beschriftet. Begründen Sie dies mithilfe des PACT-Frameworks.

### Übung

Computer und Smartphones werden immer mehr in Schulen und anderen Bildungseinrichtungen eingesetzt. Stellen Sie diese Entwicklung kurz im Bezug auf das PACT-Framework da. Welche positiven oder negativen Auswirkungen gibt es?

## 1.2 Ursprünge der MCI

Traditionell unterscheidet man drei *Generationen der Informatik*. Diese sind:

### **Erste Generation** Hardware (Rechenmaschinen)

An den Anfängen der Informatik um 1940 stehen die großen Rechenmaschinen. Diese waren so groß, dass sie einen ganzen Raum ausfüllten und waren sehr schwer zu bedienen. Es brauchte damals ein ganzes Team, um einen dieser Computer zu betreiben.

### **Zweite Generation** Software (Programmierung)

Bis zur Mitte der 1960er konnte die Computertechnik so weit verkleinert werden, dass ein Computer bequem auf einem Schreibtisch stehen konnte. So war es möglich, einen privaten Computer (PC) zu besitzen. Außerdem wurde das Berufsfeld der Informatik und der Programmierung (Software) dadurch immer stärker von der technischen Entwicklung der Hardware losgelöst.

### Dritte Generation Benutzer (Mensch-Computer-Interaktion)

Seit den 1980ern befinden wir uns in der dritten Generation der Informatik. Hier steht zunehmend der Benutzer und die Ästhetik der Hard- und Software im Mittelpunkt. Auch die Ansprüche an Computer veränderten sich. Sie wurden von komplizierten, grauen Büromaschinen zu Alltagsgegenständen und Mode-Accessoires.

In Kapitel 6 wird noch genauer auf die historische Entwicklung der Computertechnik und der Informatik eingegangen.

Neben der technischen Entwicklung hat sich auch in der Gesellschaft das Bild von Computern mit der Zeit sehr verändert. Während Computer vor 50 Jahren noch reine Forschungsmaschinen waren, ist es heutzutage nicht ungewöhnlich, Duzende von Computern in einem Haushalt zu finden: Fernseher, Radios, DVD-Player und Uhren arbeiten inzwischen alle digital und haben eigene kleine Computer eingebaut. Man spricht hier von Eingebetteten Systemen (*engl. Embedded Systems*). Derselbe technische Fortschritt macht es auch Möglich, Computer in Form von Smartphones oder Smartwatches mit sich herumzutragen. Das nennt man Mobile Computing bzw. Ubiquitous Computing (Abbildung 1.2).

Eingebettete Systeme  
Mobile Computing  
Ubiquitous Computing

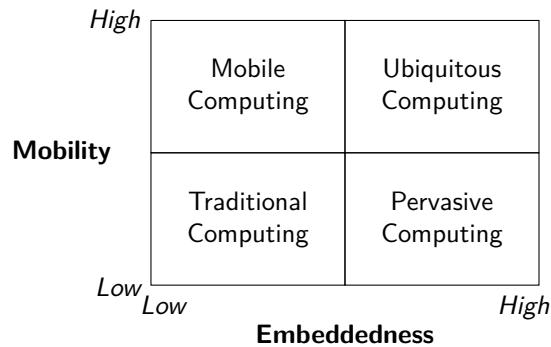


Abbildung 1.2: Ubiquitous Computing

Die Mensch-Computer-Interaktion spielt auch in der Fiktion in Filmen und Büchern eine wichtige Rolle. Technologien, die noch vor wenigen Jahren nur in Filmen zu sehen waren, gibt es heute schon zu kaufen. Besonders populär sind zum Beispiel Virtual Reality oder Smartwatches. Aber auch alternative Eingabemethoden (z.B. Sprachsteuerung oder Steuerung durch Augenbewegung) sind aktuell beliebte Forschungsgebiete der MCI. Damit wird der Trend zu höherer Mobilität und höherer Embeddedness, der schon in den vergangenen Jahrzehnten zu beobachten war, mit großer Wahrscheinlichkeit auch in der absehbaren Zukunft anhalten.

Virtual Reality

Vor diesem Hintergrund lässt sich eine erste Definition der Mensch-Computer-Interaktion aufstellen:

Mensch-Computer-Interaktion

**Definition: Mensch-Computer-Interaktion**



*„Mensch-Computer-Interaktion (MCI) (engl. Human-Computer Interaction (HCI)) als Teilgebiet der Informatik beschäftigt sich mit benutzergerechter Gestaltung von allgegenwärtigen interaktiven Systemen und ihren Mensch-Maschine-Schnittstellen.“*

— M. Dahm

### Definition: Mensch-Computer-Interaktion (2)

*„Human-computer-interaction (HCI) is a discipline concerned with the design, evaluation and implementation of ubiquitous interactive computing systems for human use and with the study of major phenomena surrounding them.“*

— Hewlett et al.

## 1.3 Warum HCI?

Nach der Beobachtung, wie sich die HCI in den letzten Jahrzehnten entwickelt hat, stellt sich nun die Frage, wieso dieses Forschungsgebiet eigentlich so wichtig ist. Am besten sieht kann man dies nachvollziehen, wenn man sich anguckt, was passiert, wenn eine Software oder ein System *nicht* benutzerfreundlich gestaltet ist:

- Ein Nutzer, der mit einem benutzerunfreundlichen System arbeiten muss, ist schnell **frustriert**.
- Besonders bei Websites wenden sich Kunden schnell anderen Angeboten zu, wenn das Design der Website nicht zufriedenstellend ist. Somit beeinflusst die Benutzerfreundlichkeit direkt den **Umsatz** eines Unternehmens.
- Schlechtes Softwaredesign führt häufig zu einem **Imageverlust** des Unternehmens, das die Software entwickelt hat.
- ...

Eine schlecht gestaltete Software kann allerdings auch noch weitaus fatalere Auswirkungen haben. In kritischen Systemen (z.B. Flugsteuerung bei Flugzeugen oder Waffensysteme) ist es essentiell wichtig, dass keine Missverständnisse bei der Benutzung des Systems auftreten, da sonst Menschenleben in Gefahr sind.

### Beispiel

Im Jahr 1990 stürzte der Flug 605 der Indian Airlines aufgrund von „Menschlichem Versagen“ ab. Beim Landeanflug in Bengaluru hatte der Pilot das Flugzeug versehentlich in einen falschen Flugmodus gebracht. Dadurch hat das Flugzeug, das sich eigentlich in der Landephase hätte befinden sollen, beschleunigt und ist in einem Golfclub abgestürzt. Es kamen 92 Menschen ums Leben.

Das Problem bei diesem Unfall lag allerdings nicht nur darin, dass der Pilot das Flugzeug in den falschen Modus gebracht hat. Scheinbar hat das Steuerungssystem des Flugzeugs dem Piloten kein ausreichendes Feedback gegeben. Hier hätte also ein besseres Design des Steuerungssystems dabei helfen können, den Unfall zu vermeiden.

Die Wichtigkeit von gutem Softwaredesign ist auch im Management von IT-Projekten bekannt. Eine Befragung von IT-Managern hat ergeben, dass der Erfolg oder Misserfolg von großen Projekten maßgeblich davon abhängt, wie sehr der Endnutzer in das Projekt involviert ist. Hier sollte man beachten, dass die Benutzerinteraktion schon in einer sehr frühen Entwicklungsphase relevant wird, nämlich bereits bei der Planung und Spezifikation der Aufgabenbereiche der Software. Lässt man den Nutzer in dieser Phase außen vor, passiert es sehr leicht, dass das Produkt am Ende gar nicht das tut, was der Benutzer erwartet. Bemerkenswert ist, dass diese Studie ergeben hat, dass das *User Involvement* sogar wichtiger ist, als *Proper Planning*, also der sorgfältigen Planung von Projekten.

*„Creating an interface is much like building a house: If you don't get the foundations right, no amount of decorating can fix the resulting structure.“*

— J. Raskin

#### ✓ Übung: Benutzerfreundliche Software

Überlegen Sie, in welchen Bereichen neben der Informatik Benutzerfreundlichkeit ein wichtiger Faktor ist. Haben Sie schon mal eine besonders benutzer(un)freundliche Software verwendet? Was hat Ihnen dabei besonders gefallen bzw. missfallen?

## 2 Menschliche Informationsverarbeitung

Um die Grundlagen der Mensch-Computer-Interaktion zu verstehen, betrachten wir in den nächsten Kapiteln zunächst das System Mensch. In diesem Kapitel geht es um den allgemeinen Prozess der menschlichen Informationsverarbeitung. Wir werden sehen, dass man dazu drei Subsysteme betrachten muss: Die Perzeption, Kognition und Motorik. Diese werden in den Kapiteln 3, 4 und ?? eingehend betrachtet. Das Ziel dieses Kapitels ist, ein Verständnis von der Mensch-Computer-Interaktion auf einem sehr groben Level zu vermitteln.

### 2.1 Benutzerschnittstellen

Die Benutzerschnittstelle ist einer der Wichtigsten Gegenstände der Mensch-Computer-Interaktion. Der Begriff wird sehr häufig zur Beschreibung von Fenstern oder Fensterinhalten auf einem Computerbildschirm benutzt. Das ist eigentlich nicht ganz richtig. Im Sinne der MCI wollen wir die folgende Definition verwenden:

#### Definition: Benutzerschnittstelle

Jede Interaktion zwischen Mensch und Computer erfolgt über eine Benutzerschnittstelle (*engl. User Interface, UI*). Die Benutzerschnittstelle ermöglicht eine 2-Wege-Kommunikation:

- Das Senden von Steuersignalen **vom Menschen zum Computer** über **Eingabegeräte** (z.B. Maus, Tastatur, Joystick, Spracheingabe, ...)
- Das Präsentieren von Ergebnissen **vom Computer zum Menschen** über **Ausgabegeräte** (z.B. Monitor, Lautsprecher, Haptisches Feedback, ...)

Eingabegerät

Ausgabegerät

Der Begriff Benutzerschnittstelle ist unabhängig davon, ob die Interaktion über visuelle Informationen oder zum Beispiel Sprache oder Gestik erfolgt.

Auf Grundlage dieser Definition, lässt sich bereits ein sehr abstraktes Interaktionsframework aufbauen, das die Interaktion zwischen Benutzer und Computer modelliert (Abbildung 2.1).

Interaktionsframework

Die Abbildung zeigt, dass die Interaktion aus vier wesentlichen Aktionen besteht:

1. Eine **Ausgabe** des Systems wird auf der UI „angezeigt“<sup>1</sup>. Der Benutzer macht eine entsprechende **Beobachtung**.

<sup>1</sup>Je nach Schnittstelle kann dies jede Sorte Feedback wie zum Beispiel Bildschirmausgaben, Tonausgaben oder haptisches Feedback bedeuten.

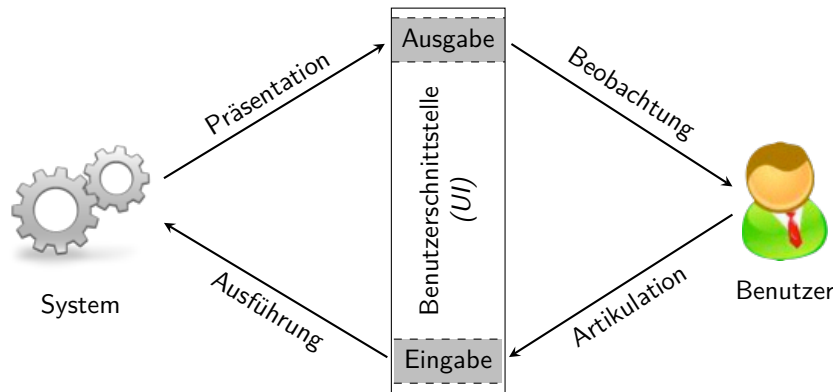


Abbildung 2.1: Interaktionsframework

Interaktionsframework

2. Basierend auf der Beobachtung führt der Benutzer eine Aktion (**Eingabe**) aus. Diese Aktion nennt man **Artikulation**.
3. Im System wird ein Arbeitsvorgang **ausgeführt**.
4. Das System **präsentiert** ein Ergebnis auf der UI. Diese **Ausgabe** kann wieder vom Nutzer beobachtet werden, wodurch der Kreislauf von vorne beginnt.

Am Interaktionsframework kann man schon sehen, dass die Benutzerschnittstelle eine sehr zentrale Rolle in der Mensch-Computer-Interaktion einnimmt. Bevor wir uns damit in Kapitel 7 eingehend beschäftigen, betrachten wir jedoch zunächst den Menschen und den Computer getrennt voneinander.

## 2.2 Informationspsychologie

*„The domain of concern to us [...] is how humans interact with computers. A scientific psychology should help us in arranging the interface so it is easy, efficient and error free—even enjoyable.“*

— S. Card, T. Moran, A. Newell: Model Human Processor, 1983

Der Mensch ist ein vielseitiges und sehr komplexes Lebewesen. Es ist nicht möglich, alle Facetten gleichzeitig zu betrachten und eine Art *allumfassendes* Modell zu entwickeln. Stattdessen werden wir nur ausgewählte Aspekte zur Zeit betrachten und dieser Grundlage Modelle für die menschliche Informationsverarbeitung entwickeln. Dies fällt in das Forschungsgebiet der Informationspsychologie.

Intuitiv fallen schon mehrere Sichten auf den Menschen im Sinne der MCI ein:

### Perzeptuelle Sicht

Der Mensch ist ein System, das Informationen über seine Sinne (sehen, hören, ...) aufnehmen kann.

### **Konzeptionelle Sicht Informationsverarbeiter**

Der Mensch ist ein System, das aufgenommene Informationen verarbeitet und in einen gemeinsamen Kontext setzt. Also Informationsverarbeiter kann der Mensch Rückmeldungen von (Computer-)Systemen verstehen und entsprechende weitere Arbeitsschritte planen.

### **Physikalische Sicht**

Der Mensch wird hier als Akteur gesehen, der z.B. Knöpfe drücken kann, die Maus bewegt usw. und so eine Eingabe für den Computer produziert. In der Forschung über Körper und Ergonomie (*engl. Human Factors*) betrachten wir den Menschen meistens aus der physikalischen Sicht.

### **Verarbeiter natürlicher Sprache und Kommunikation**

Bei dieser Sicht auf den Menschen wendet man Erkenntnisse aus dem Forschungsgebiet der Linguistik an. Technologien wie Siri arbeiten sehr stark mit dem Menschen als Verarbeiter natürlicher Sprache.

### **Informavore und soziales Wesen**

Mit dieser Sichtweise beschäftigen sich Wissenschaften wie die Soziologie, Anthropologie oder die Informations- und Kommunikationswissenschaft. Im Kontext der MCI spielt sie vor allem bei Technologien wie sozialen Netzwerken eine Rolle.

In Folgenden versuchen wir ein Modell für den Menschen als Informationsverarbeiter zu finden, das sich gut für die Anwendung in der Mensch-Computer-Interaktion eignet. Wir betrachten den Menschen dazu zunächst aus der perceptuellen und physikalischen Sicht (Abbildung 2.2).

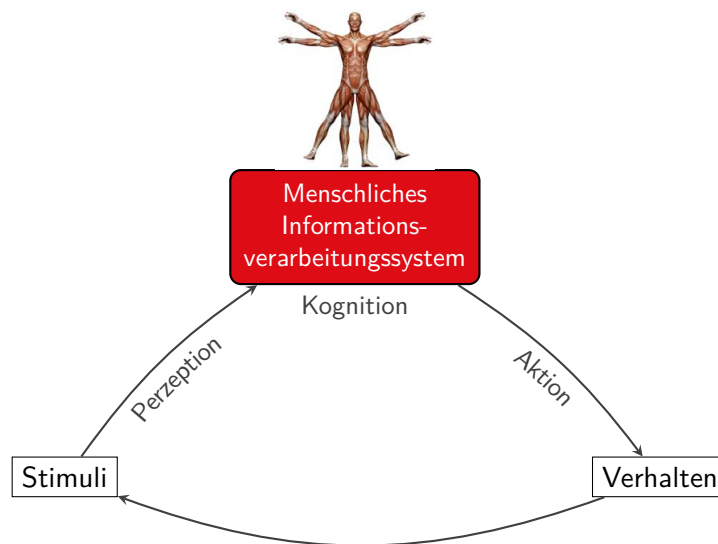


Abbildung 2.2: Menschliche Informationsverarbeitung

Wie bereits im Interaktionsframework (Abbildung 2.1) beim „System“ modellieren wir den Menschen als Informationsverarbeitungssystem mit Eingabe, Verarbeitung und Ausgabe. Die Eingabe (Stimuli) gelangt über die Perzeption (Wahrnehmung) in das System Mensch hinein und wird dort durch die Kognition. Basierend auf den Ergebnissen der Kognition wird eine entsprechende Ausgabe (Aktion) durch die Motorik ausgeführt. Da jede Ausgabe die Umgebung des Menschen verändern kann, führt eine Aktion normalerweise zu einem erneuten Reiz, der wiederum durch die Perzeption von Menschen wahrgenommen wird. Dadurch entsteht ein kontinuierlicher Kreislauf von Eingabe und Ausgabe. Die einzelnen Komponenten der Abbildung werden in den Kapiteln 3, 4 und ?? eingehend betrachtet.

Dieses Modell kann nun ein wenig verfeinert werden, indem man die Menschliche Informationsverarbeitung in der Biologie betrachtet (Abbildung 2.3).

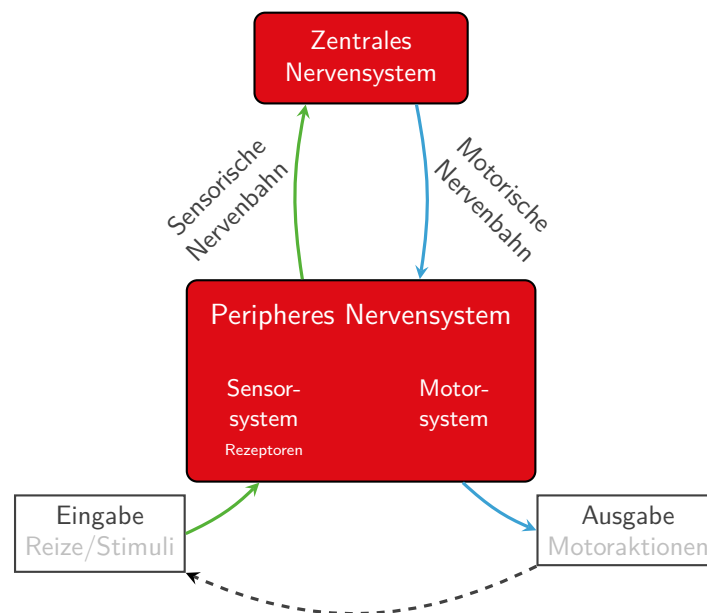


Abbildung 2.3: Mensch I/O

Die Verarbeitung von Informationen läuft nach folgenden Schritten ab:

1. Die **Eingabe** wird durch die Sinne des Menschen wahrgenommen.
2. Spezielle **Rezeptoren** wandeln die Daten der Sinnesorgane in elektrische Impulse um.
3. Die elektrischen Impulse werden über die **sensorischen Nervenbahnen** im peripheren Nervensystem zum **zentralen Nervensystem** (Gehirn) geleitet.
4. Das Gehirn verarbeitet die Information.

5. Als Reaktion sendet das Gehirn elektrische Signale über die **motorischen Nervenbahnen** zurück über das periphere Nervensystem an spezielle **motorische Neuronen**.
6. Die motorischen Neuronen lösen als **Ausgabe** Muskelbewegungen aus.

Im Sinne der Informationspsychologie ist also Wesen bzw. System mit drei Kernkompetenzen:

### 1. Informationsaufnahme *engl. Perception*

Die Informationsaufnahme ist der Teil des menschlichen Systems, der dafür zuständig ist, Reize aus der Umwelt in Signale umzuwandeln, die von Gehirn weiterverarbeitet werden können. Solche Reize werden durch die verschiedenen Sinnesorgane aufgenommen (man spricht hier von **sensorischen Modalitäten**, siehe Tabelle 2.1). Für die Mensch-Computer-Interaktion sind aber besonders der visuelle Kanal (sehen), der auditive Kanal (hören) und der haptische Kanal (fühlen) wichtig. Im Detail wird hierauf noch in Kapitel 3 eingegangen.

Modalität

Wahrnehmungssystem	Organ	Empfindung
Visuell	Auge	Farbe, Helligkeit
Auditiv	Ohr	Tonhöhe, Lautstärke
Haptisch	Haut, Muskelspindel	Vibration, Druck, Stellung
Thermisch	Haut	Wärme, Kälte
Olfaktorisch	Nase	Geruch
Gustatorisch	Mund	Geschmack
...	...	...

Tabelle 2.1: Sensorische Modalitäten

### 2. Informationstransformation und -speicherung *engl. Cognition*

Die Informationsspeicherung findet größtenteils im Gehirn statt. Man unterscheidet

- Sensorische Speicher (Ultrakurzzeitgedächtnis)
- Arbeitsgedächtnis (Kurzzeitgedächtnis)
- Langzeitgedächtnis

Außerdem kann das Gehirn Informationen verarbeiten und daraus Schlüsse ziehen. Man kann grob drei Arten von Informationstransformationen unterscheiden:

- Logisches Denken
- Problemlösung
- Aneignung von Fähigkeiten

Sowohl mit Informationsspeicherung als auch mit der -transformation beschäftigen wir uns in Kapitel 4.

### 3. Motorische Aktionen *engl. Motor Action*

Am Ende der Informationsverarbeitungskette steht die Ausführung von Aktionen. Dafür sind sogenannte **Endeffektoren** zuständig. Dazu zählen u.A. Arme, Beine, Augen, Stimmbänder oder die Körperhaltung. Hiermit beschäftigen wir uns in Kapitel ??.

## 2.3 Modelle der menschlichen Informationsverarbeitung

### 2.3.1 Model Human Processor

Auf der Grundlage des vorhergegangenen Abschnittes können wir jetzt ein Modell des Menschen für die Mensch-Computer-Interaktion definieren: Das Model Human Processor. Es gibt auch noch andere Modelle, die in Abschnitt 2.3.2 und 2.3.3 kurz erläutert werden. Danach werden wir aber nur noch mit dem Model Human Processor arbeiten.

Model Human  
Processor

#### Definition: Model Human Processor

Das Model Human Processor wurde am Xerox Parc entwickelt. Es modelliert den Menschen als ein System, das aus drei Subsystemen zusammengesetzt ist:

1. Das **Perzeptuelle (Eingabe-)system** nimmt Stimuli aus der Umwelt auf.
2. Ein **Kognitives System** zur rationalen Verarbeitung der Informationen.
3. Ein **Motor (Ausgabe-)system** zur Kontrolle von Aktionen.

Jede Information, die vom Menschen verarbeitet wird, durchläuft die drei Subsysteme in diesem Modell sequentiell hintereinander. Analog zu traditionellen Informationssystemen, spricht man davon, dass jedes Subsystem einen eigenen **Prozessor** und **Speicher** hat. Empirische Untersuchungen haben die folgenden Laufzeiten für die einzelnen Prozessoren ergeben:

1. **Perzeptueller Prozessor:**  $\tau_P = 100[50 \sim 200]$ ms
2. **Kognitiver Prozessor:**  $\tau_C = 100[25 \sim 170]$ ms
3. **Motorischer Prozessor:**  $\tau_M = 100[70 \sim 360]$ ms

*Hinweis zum Lesen: Angegeben sind immer die mittlere, kürzeste und längste Zeit für den Prozessor. Meistens beschränken wir uns auf die mittlere Zeit, also ungefähr 100 ms pro Prozessor.*

#### Aufgabe 2: Bilder pro Sekunde

Wie viele Bilder pro Sekunde müssen BetrachterInnen mindestens gezeigt werden,



damit die Illusion einer Bewegung entsteht? Nutzen Sie das Model Human Processor für Ihre Argumentation.

**Filterung und Selektion** Über die vielen Sinnesorgane nimmt der Mensch kontinuierlich Daten auf. D. McCandless hat für die unterschiedlichen Sinne folgende Datenraten abgeschätzt:

**Sehen:** 1250 MB/s

**Fühlen:** 125 MB/s

**Hören/Riechen:** 12,5 MB/s

**Schmecken:** 1,25 MB/s

**Andere:** < 1 MB/s

Für das Gehirn sind das zu viele Daten, um alles speichern oder verarbeiten zu können. Daher müssen diese Daten zuerst gefiltert und selektiert werden. Die Filterung passiert bereits bei der Perzeption. Die Sinnesorgane selbst filtern automatisch so viele Daten, dass von den insgesamt 1,5 Gbit/s nur ca. 15 Mbit/s bei den Rezeptoren ankommen.

Durch die darauf folgende Selektion wird diese Datenmenge noch einmal auf 100 bit/s verkleinert, die tatsächlich das Bewusstsein erreichen. Die Grenzen für die verarbeitbaren Datenmengen hängen von vielen Faktoren ab und variieren von Mensch zu Mensch sehr stark, sodass man keine allgemeinen Grenzwerte angeben kann.

Es ist wichtig zwischen Filterung und Selektion zu unterscheiden:

- Filterung passiert automatisch und kann auch nicht beeinflusst werden. Meistens hängt die Filterung von den physischen Fähigkeiten der Sinnesorgane des Menschen ab. In der visuellen Wahrnehmung wird zum Beispiel dadurch gefiltert, dass das menschlich Auge gar nicht das gesamte Sichtfeld scharf sehen kann.
- Selektion ist ein semi-aktiver Prozess, der meistens bewusst gesteuert wird. Die Entscheidung, welche Informationen gefiltert werden, hängt davon ab, worauf man seine Aufmerksamkeit richtet.

Das Model Human Processor ist in seiner Gesamtheit in Abbildung 2.4 dargestellt.

### 2.3.2 Soar State, Operator, Apply, Result

Soar (ursprünglich Akronym für State, Operator, Apply, Result) ist ein weiteres Modell, das den Prozess der menschlichen Informationsverarbeitung beschreiben soll. Bei Soar definiert man einige **primitiven Prinzipien** als Grundlage der menschlichen Kognition. Besonders nützlich ist dieses Modell bei der Entwicklung von künstlicher Intelligenz, da man hier dieselben primitiven Prinzipien einpflegen kann, um so menschliche Verhaltensweisen zu simulieren. Die primitiven Prinzipien sind:

1. Problemlösen wird als Suche in Problemräumen repräsentiert.

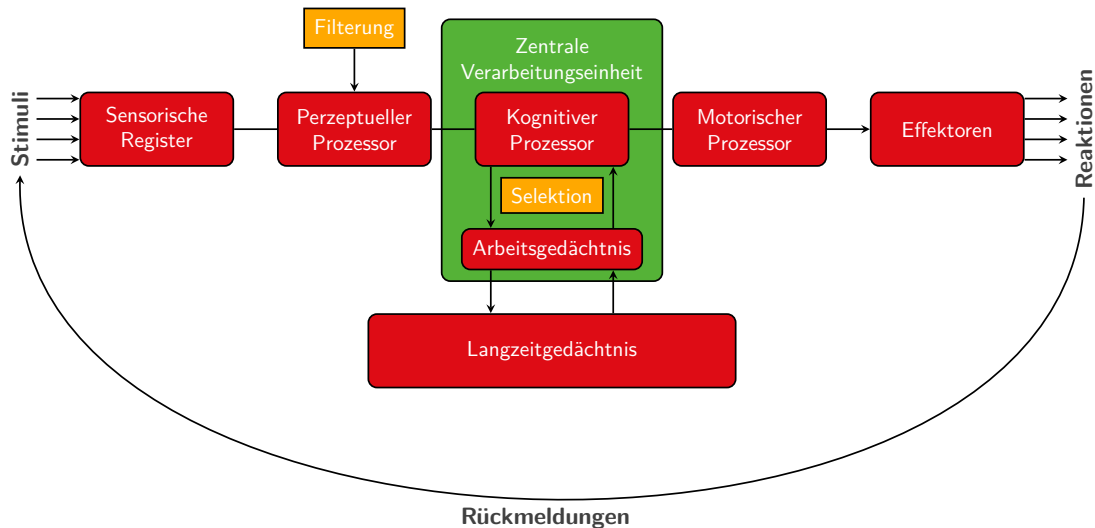


Abbildung 2.4: Model Human Processor

2. Dauerhaftes Wissen wird durch Produktionsregeld im Langzeitgedächtnis repräsentiert.
3. Temporäres Wissen wird durch Objekte im Arbeitsspeicher repräsentiert.
4. Neue Ziele werden nur generiert, wenn Sackgassen (*engl. Impasses*) auftreten.
5. Lernmechanismus: *Chunking and Reinforcement Learning*.

### 2.3.3 Weitere Modelle

Es gibt noch weitere Modelle zur menschlichen Informationsverarbeitung, wie zum Beispiel das EPIC-Modell von D.E. Kieras. In den meisten Modellen findet man aber die drei Kernkomponenten des Model Human Processors wieder (Perzeption, Kognition, Motorik). Im folgenden wird nur noch das Model Human Processor verwendet.

## 2.4 Prättentive und attentive Merkmale

Der Mensch ist ein sehr komplexes System. Bei der Betrachtung von interaktiven Systemen ist auch in der Regel die Betrachtung des Menschen und nicht die Betrachtung des Computersystems der limitierende Faktor. Deshalb vereinfacht man das System Mensch meistens sehr stark. Das ist auch beim Model Human Processor der Fall. Im Laufe dieser Veranstaltung wird diese Vereinfachung allerdings an verschiedenen Stellen etwas gelockert werden. Eine dieser Stellen ist der Übergang zwischen Perzeption und Kognition:

Model Human Processor

Im Model Human Processor sind die drei Subsysteme Perzeption, Kognition und Motorik klar voneinander getrennt. In der Realität ist das jedoch nicht der Fall. Das zeigt das folgende Beispiel:

28348682794629478392	28 <b>3</b> 48682794629478 <b>3</b> 92
7490709237982359812	74907092 <b>3</b> 7982 <b>3</b> 59812
13496198346198264959	<b>1</b> 3496198 <b>3</b> 46198264959
0739087392807916212	07 <b>3</b> 9087 <b>3</b> 92807916212
73498398983247829347	<b>7</b> 3498 <b>3</b> 9898 <b>3</b> 247829 <b>3</b> 47
9817349817498172324	9817 <b>3</b> 49817498172 <b>3</b> 24
<b>Attentive Aufgabe</b>	<b>Präattentive Aufgabe</b>

Abbildung 2.5: Attentive und Präattentive Aufgabe

**Die Aufgabe** Zählen Sei die Dreien.

Nach dem Model Human Processor bräuchte man für beide abgebildeten Situationen die gleiche Zeit. Man muss für jede Ziffer mindestens den perzeptuellen und den kognitiven Prozessor durchlaufen lassen, um die Ziffern zu erkennen und mit der Aufgabenstellung zu vergleichen („Ist es eine 3?“). Insgesamt sind es 117 Ziffern, also kann man mit einer durchschnittlichen Bearbeitungszeit von  $117 \cdot (\tau_P + \tau_C) = 117 \cdot 200 \text{ ms} = 23,4 \text{ s}$  rechnen. In der ersten Situation mag das stimmen, in der zweiten Situation stimmt diese Zeit aber offensichtlich nicht.

#### Definition: Präattentive und Attentive Merkmale

Es gibt bestimmte Merkmale, die bereits vor dem kognitiven System erkannt werden. Diese nennt man **Präattentive Merkmale**. Dazu zählen unter Anderem Form, Farbe, Geschlossenheit und Bewegung. Alle anderen Merkmale, die erst im kognitiven System verarbeitet werden, nennt man **Attentive Merkmale**. Die Unterschiede von attentiven und präattentive Merkmalen sind in Tabelle 2.2 aufgelistet.

Präattentive  
Merkmale

Attentive Merkmale

Attentiv	Präattentiv
Kontrolliert	Automatisch
Benutzt zumeist Fovea	Auch peripher möglich
Detailliert	Oberflächlich
Seriell	Parallel
Lamgsam	Schnell
Kann unterdrückt werden	Nicht zu unterdrücken
Bewusstseinspflichtig	Unbewusst
Beansprucht Aufmerksamkeit	Voraufmerksam

Tabelle 2.2: Attentive und Präattentive Merkmale

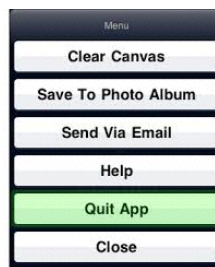
Präattentive Aufgaben können in konstanter Zeit gelöst werden. Wie lange das genau ist, hängt vom Individuum ab, allerdings lässt sich die Bearbeitungszeit auf ungefähr 200 ms abschätzen. Im Vergleich zu einem einfachen attentiven Problem ist dies noch kein großer Unterschied, allerdings skalieren präattentive Aufgaben sehr gut. Zum Beispiel würde in Abbildung 2.5 das hinzufügen von Ziffern, die keine 3 sind, die attentive Aufgabe deutlich schwieriger machen. Im zweiten Fall würde sich die Bearbeitungszeit aber nicht ändern. Mit dieser neuen Definition von präattentiven Merkmalen lässt sich die Bearbeitungszeit für die obige Aufgabe neu abschätzen (siehe Aufgabe).

### Aufgabe 3: Präattentive Verarbeitungszeit

Wenden Sie das Model Human Processor und Ihr Wissen über attentive präattentive Merkmale an, um abzuschätzen, wie lange das **Zählen der roten Dreien** in Abbildung 2.5 dauert.

### Aufgabe 4: Präattentive Merkmale

Wie wurden in der folgenden Abbildung präattentive Merkmale zur Gestaltung des User Interfaces benutzt?



## 3 Wahrnehmung

In diesem Kapitel geht es um die menschliche Wahrnehmung. Im Model Human Processor wird hier also der perzeptuelle Prozessor näher betrachtet. Für die Mensch-Computer-Interaktion ist es essentiell zu verstehen, wie die menschliche Wahrnehmung funktioniert, damit man Computersysteme konstruieren kann, die vom Menschen gut wahrgenommen werden können. Besonders wichtig ist auch zu verstehen, wo die Grenzen der menschlichen Wahrnehmung sind und welche Arten der Informationsvermittlung man bei einem Informationssystem vermeiden sollte.

Im Anwendungsbereich der Mensch-Computer-Interaktion spielen fast ausschließlich die visuelle und die auditive Wahrnehmung eine Rolle. Das hat vor allem den Grund, dass der Mensch durch das Sehen bzw. Hören die meisten Informationen aufnimmt und diese am besten verarbeiten kann. In der Forschung gibt es auch Ansätze die anderen Sinne des Menschen auszunutzen, allerdings sind die meisten dieser Geräte nicht besonders alltagstauglich.

### 3.1 Wahrnehmungstheorie

In diesem Abschnitt werden wir einige Eigenschaften der menschlichen Wahrnehmung kennen lernen, die über alle Modalitäten hinweg gleich sind. Dies wird die Grundlage und Motivation dafür sein, die einzelnen Wahrnehmungskanäle des Menschen genauer zu betrachten.

Modalität

#### Definition: Wahrnehmungstheorie

Die Wahrnehmungstheorie (abzugrenzen von der Psychophysik, siehe Abschnitt 3.7) beschäftigt sich mit der Kluft zwischen subjektiv-psychologischem Erleben bei Wahrnehmungen und objektiv-physiologischer Messbarkeit der Vorgänge.

In der Wahrnehmungstheorie wird also die Verbindung zwischen umgangssprachlichen Ausdrücken wie „Das ist mir zu laut!“ und den damit zusammenhängenden messbaren physikalischen Phänomenen hergestellt. Wir betrachten dafür die Wahrnehmung des Menschen zunächst nur auf einem qualitativen Level. In der Psychophysik (Abschnitt 3.7) wird es dann darum gehen, die subjektiv wahrgenommenen Ereignisse durch objektiv messbare Werte zu belegen.

Für die Mensch-Computer-Interaktion sind die Erkenntnisse aus der Wahrnehmungstheorie sehr wichtig, wenn es darum geht, Signale an den Benutzer zu senden. So kann man beispielsweise sicherstellen, dass ein Warnsignal definitiv laut genug ist, dass der Benutzer dieses auch hören kann. Genauso findet die Wahrnehmungstheorie beim Design

der Ausgabegeräte des Computers anwendung (Bildschirmhelligkeit, Kontraste, Lautstärke etc.).

### Definition: **Wahrnehmung**

Wahrnehmung im Sinne der Wahrnehmungstheorie ist ein Prozess und das Ergebnis der Informationsverarbeitung von Reizen aus der Umwelt (Exterozeption) und dem Körperinneren (Interozeption).

Exterozeption  
Interozeption

Die obige Definition passt nicht ganz zu der umgangssprachlichen Wahrnehmung. Wenn wir umgangssprachlich von Wahrnehmung reden, beziehen wir uns häufig auf Objekte oder Kombinationen von Sinneseindrücken. Im Sinne der Wahrnehmungstheorie beinhaltet das jedoch schon zusätzliche Verarbeitungsschritte:

**Sensorische Integration** Verschiedene Sinneseindrücke (Reize), die in einem Zeitraum aufgetreten sind, werden nach verschiedenen Wahrscheinlichkeiten gewichtet und zusammengeführt. Mehr dazu in Abschnitt 3.7.

Reiz

**Interpretation** Der kombinierte Eindruck durch die Reize wird dann noch mit Informationen aus dem Gedächtnis kombiniert. Dabei werden fehlende Informationen probabilistisch ergänzt (sofern das denn möglich ist).

Daraus geht hervor, dass die Wahrgenommene Umwelt *nie* eine direkte Abbildung der physikalischen Realität ist, sondern immer eine Interpretation auf Grundlage der wahrgenommenen Reize (Perzepte).

Perzept

### **Aufgabe 5: Wahrnehmungstheorie**

Sie sitzen in einem Flugzeug und hören eine Sirene. In welchem Schritt der umgangssprachlichen Wahrnehmung entscheiden Sie (unterbewusst), dass es sich hier vermutlich nicht um ein Polizeiauto handelt?

## **Stufenmodell der Wahrnehmung**

Wie wir oben bereits gesehen haben, hat der Prozess der Wahrnehmung sequentiellen Charakter. Nach einem Stimulus werden also nacheinander verschiedene Arbeitsabläufe ausgeführt, die im Prinzip unabhängig voneinander sind. Das motiviert das Stufenmodell der Wahrnehmung (Abbildung 3.1).

Stimulus

Es gibt nun zwei Interpretationen dieses Modells:

### **1. Bottom-Up Data Driven**

Ein Stimulus ist immer das erste Ereignis im Modell. Basierend auf dem Reiz werden verschiedene Arbeitsschritte ausgeführt, um den Reiz zu interpretieren und in einen Kontext einzuordnen. Diese Theorie haben wir auch beim Model Human Processor angewendet.

Model Human  
Processor

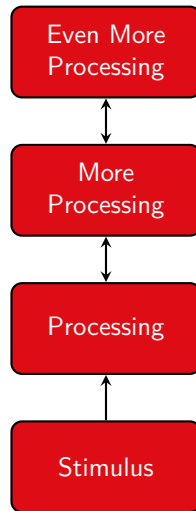


Abbildung 3.1: Stufenmodell

## 2. Top-Down Hypothesis Driven

Verschiedene Forschungen haben ergeben, dass der sequentielle Ablauf von unten nach oben nicht immer stimmt. Tatsächlich ist es so, dass die Ergebnisse von höheren Arbeitsschritten die Arbeitsweise der Schritte darunter direkt beeinflussen können. Das zeigt das folgende Beispiel (Abbildung 3.2):



Abbildung 3.2: Top-Down Processing

Je nachdem, ob man das Bild von links nach rechts oder von oben nach unten liest, erkennt man in der Mitte den Buchstaben B oder die Zahl 13. Offensichtlich beeinflusst hier also vorhergegangenes Wissen die Art und Weise, wie man eine Information wahrnimmt.

Heutzutage ist man sich relativ einig, dass sowohl das Bottom-Up Processing als auch das Top-Down Processing beim Menschen tatsächlich zum Einsatz kommen.

## 3.2 Physikalische Grundlagen

**Analoge Signale** Bevor wir die einzelnen Wahrnehmungskanäle des Menschen betrachten, müssen wir uns klar machen, dass es eine starke Diskrepanz zwischen Computersignalen und „Menschensignalen“ gibt, die überbrückt werden muss. Computer arbeiten mit digitalen Signalen. Physikalisch gesehen handelt es sich dabei um *diskrete Werte* zwischen denen sprunghaft gewechselt wird. Erhöht man zum Beispiel die Lautstärke, wird im Computer irgendwo eine Zahl auf einen bestimmten Wert *gesetzt*. Der Übergang zwischen der alten und der neuen Lautstärke erfolgt sofort und ohne Zwischenstufen. Bei modernen Computern ist die Abstufung aber meistens so klein, dass man keinen Unterschied wahrnimmt.

In der Welt außerhalb des Computers werden die meisten Informationen hingegen über analoge Signale übertragen. Analoge Signale definieren einen *kontinuierlichen Verlauf* einer physikalischen Größe. Hier gibt es sozusagen unendlich viele Zwischenstufen zwischen zwei Werten. Bei der Interaktion zwischen Mensch und Computer müssen jetzt bei der Ein- und Ausgabe digitale und analoge Signale ineinander umgewandelt werden. Dafür ist die Hardware des Computers zuständig. Für auditive Signale gibt es zum Beispiel einen Digital-Analog-Wandler im Computer, der die digitalen Audiosignale in analoge Signale umwandelt. Auf der visuellen Ebene übernimmt der Bildschirm diese Aufgabe.

**Wellen** In den folgenden Abschnitten werden wir sehen, dass sehr viele analoge Informationen periodisch durch Wellen übermittelt werden. Man kann sich Wellen in etwa so vorstellen, wie die Wellen, die sich um einen ins Wasser geworfenen Stein ausbreiten. Man unterscheidet zwei Arten von Wellen:

**Transversalwellen** Diese Wellen schwingen senkrecht zur Ausbreitungsrichtung. Das ist zum Beispiel bei Wellen im Wasser der Fall.

**Longitudinalwellen** Diese Wellen schwingen in der Ausbreitungsrichtung. Das ist zum Beispiel bei Wellen in der Luft der Fall.

Jede Welle lässt sich abgesehen von der Schwingungsrichtung durch folgende Eigenschaften charakterisieren:

**Die Phase** beschreibt den sich wiederholenden Verlauf in einem festen zeitlichen Abstand.

**Die Amplitude** gibt an, wie stark die einzelnen Oszillatoren der Welle schwingen. Das entspricht dem maximalen Wert in der Phase.

**Die Wellenlänge** ist die Entfernung von zwei „Wellenbergen“.

**Die Frequenz** gibt die Anzahl von sich wiederholenden Schwingungen pro Sekunde an. Die Frequenz wird in Hz angegeben.



**Die Ausbreitungsgeschwindigkeit** gibt an, wie schnell sich die Welle ausbreitet. Diese Geschwindigkeit hängt vor allem vom Medium ab, in dem sich die Welle ausbreitet.

Wellenlänge und Frequenz lassen sich mit der Formel

$$c = \lambda \cdot f$$

ineinander umrechnen. Dabei ist  $c$  die Ausbreitungsgeschwindigkeit der Welle in m/s,  $\lambda$  die Wellenlänge in m und  $f$  die Frequenz in Hz.

### 3.3 Visuelle Wahrnehmung

Wir haben bereits in Abschnitt 2.3.1 gesehen, dass die visuelle Wahrnehmung der primäre Weg der Informationsaufnahme des Menschen ist. Entsprechend ausführlich beschäftigen wir uns hier mit der visuellen Informationsaufnahme. Dazu betrachten wir zunächst die physikalischen und biologischen Hintergründe des Sehens. Danach werden wir uns mit dem Prozess der Objekt- und Gesichtserkennung beschäftigen und daraus Konsequenzen für die Mensch-Computer-Interaktion ziehen.

#### 3.3.1 Physikalische Hintergründe

##### Was ist Licht?

Die korrekte Definition von Licht ist die folgende:

**Definition: Licht**

Licht ist der sichtbare Bereich des elektromagnetischen Spektrums. Es umfasst Wellenlängen zwischen etwa 380 nm und 750 nm.

Das gesamte elektromagnetische Spektrum enthält noch weitere Arten von Strahlung, wie zum Beispiel Infrarot- oder ultraviolette Strahlung. Es ist in Abbildung 3.3 dargestellt. Man kann Licht als elektromagnetische Welle oder als Teilchen (Menge von Photonen) betrachten. Je nachdem welches Experiment man ausführt, erhält man Ergebnisse, die für den Wellen- bzw. Teilchencharakter von Licht sprechen. In der Physik arbeitet man daher mit Licht sowohl als Welle als auch als Teilchen. Für uns reicht es Licht als Welle zu betrachten.

Als Welle hat Licht eine Amplitude und eine Frequenz. Die Frequenz des Lichtes definiert die Farbe des Lichtes, die Amplitude dessen Helligkeit.

**Die Frequenz** der Welle definiert die Farbe des Lichtes. Je niedriger die Frequenz ist, desto röter ist das Licht, je höher die Frequenz ist, desto blauer ist das Licht. Für eine gegebene Frequenz kann man die entsprechende Farbe im elektromagnetischen Spektrum nachschlagen (Abbildung 3.3).

**Die Wellenlänge** wird oft anstelle der Frequenz verwendet. Das ist auch in Abbildung 3.3 der Fall. Die Wellenlänge wird in m, bzw. für Licht meistens in nm angegeben.

**Die Ausbreitungsgeschwindigkeit** einer Lichtwelle ist die Lichtgeschwindigkeit, also  $300 \cdot 10^6$  m/s.

**Die Amplitude** definiert die Helligkeit des Lichtes.

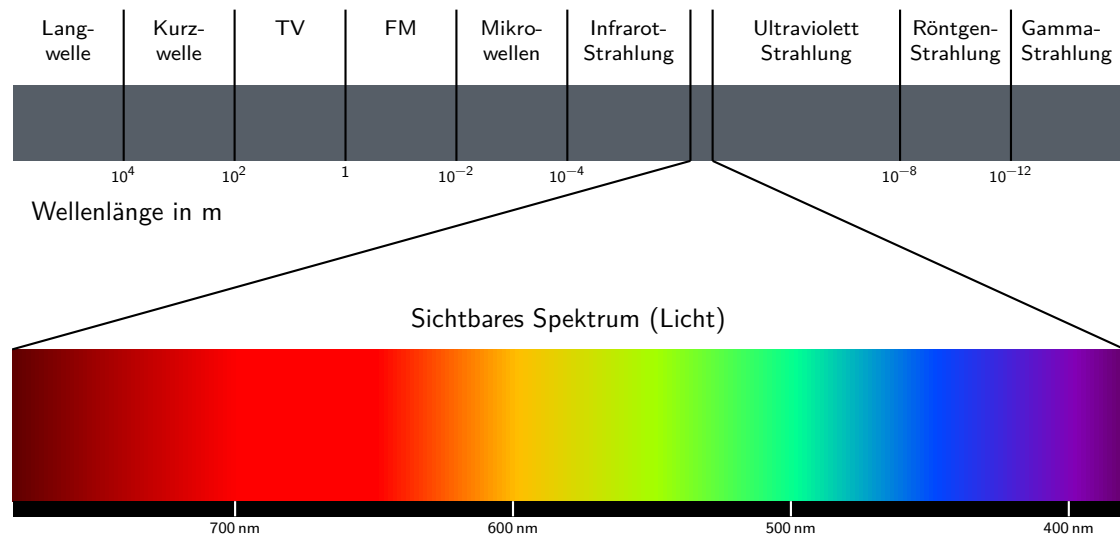


Abbildung 3.3: Elektromagnetisches Spektrum

#### Aufgabe 6: Wellenlänge und Frequenz

Rechnen Sie die folgenden Wellenlängen in Frequenzen, bzw. Frequenzen in Wellenlängen um und geben Sie die Farbe des Lichtes an:

1. 500 nm
2.  $1,4 \cdot 10^{14}$  Hz

### Absorption und Reflektion

Damit ein Objekt vom Menschen (oder zum Beispiel einer Kamera) wahrgenommen werden kann, muss Licht von einer Lichtquelle ausgesendet werden, dann auf das Objekt treffen und schließlich ins menschliche Auge gelangen. Gelangt das Licht danach nicht ins Auge, kann das Objekt auch nicht gesehen werden. Beim Auftreffen des Lichtes auf das Objekt treten zwei Effekte auf:

**Absorption** Von dem eintreffenden Licht wird ein bestimmter Anteil absorbiert. Manche Wellenlängen werden vollständig absorbiert, andere nur zu einem gewissen Anteil.

**Reflektion** Der Teil des Lichtes, der nicht absorbiert wird, wird reflektiert und kann z.B. von einer Kamera „gesehen“ werden.

Die Farbe eines Objektes hängt letztlich nur davon ab, welche Wellenlängen von dessen Oberfläche reflektiert werden. Eine Tomate reflektiert beispielsweise fast nur Wellenlängen aus dem roten Farbbereich, sodass sie uns rot erscheint. Ebenso absorbieren grüne Blätter fast alle Wellenlängen, die nicht aus dem grünen Farbbereich sind. Eine weiße Oberfläche reflektiert alle sichtbaren Wellenlängen, eine schwarze Oberfläche absorbiert alle<sup>1</sup>.

#### Aufgabe 7: Absorption und Reflektion

1. Was passiert, wenn man eine Tomate mit blauem Licht bestrahlt?
2. Wenn man nachts aus einem hell erleuchteten Raum ins Dunkle sehen will, sieht man nur sein Spiegelbild. Tagsüber passiert das nicht. Woran könnte das liegen?
3. Angenommen es ist möglich, sich unsichtbar zu machen. Welche physikalischen Probleme könnte das mit sich bringen (bezogen auf obige Informationen)?

## Optik

In der Physik unterscheidet man verschiedene Arten von Optik (z.B. Wellenoptik oder Strahlenoptik). Wir wenden hier die Strahlenoptik. Man kann die Strahlenoptik als eine Vereinfachung der Wellenoptik betrachten. Wichtig ist aber, dass die Strahlen- und Wellenoptik nicht im Konflikt zueinander stehen. Wir betrachten also auch hier das Licht als Welle.

Um zu verstehen, wie das menschliche Auge Objekte fokussiert, müssen wir die Arbeitsweise einer Linse verstehen. Eine Linse lenkt Licht auf eine bestimmte Weise um, die dazu führt, dass parallel eintreffende Lichtstrahlen in einem Punkt fokussiert werden. Dieser Punkt heißt Brennpunkt (siehe Abbildung 3.4). Den Abstand zwischen Linse und Brennpunkt bezeichnet man als Brennweite oder einfach nur  $f$ . Die Brennweite ist eine Eigenschaft der Linse.

Brennpunkt

Parallele Lichtstrahlen sind für unsere Betrachtungen eher uninteressant. Wir gehen davon aus, dass jeder Punkt eintreffendes Licht in alle Richtungen reflektiert. Es treffen also von jedem Punkt aus auch mehrere Lichtstrahlen auf die Linse (siehe Abbildung 3.5). Damit man diesen Punkt scharf auf einer Bildebene in einer gewissen Entfernung von der Linse abbilden kann, muss die Linse die entsprechende Brennweite haben. Ist dies der Fall, wird jeder Punkt eines Objektes genau auf einen Punkt auf der Bildebene abgebildet.

<sup>1</sup>In der echten Welt gibt es kein perfektes schwarz und kein perfektes weiß. Jedes Material absorbiert und reflektiert immer einen kleinen Teil des einfallenden Lichtes. Genauso ist es eher selten, dass ein Objekt genau eine Wellenlänge reflektiert. Meistens werden sehr viele Wellenlängen zu einem geringen Anteil reflektiert, aber nur eine bestimmte Farbe zu einem hohen Anteil.

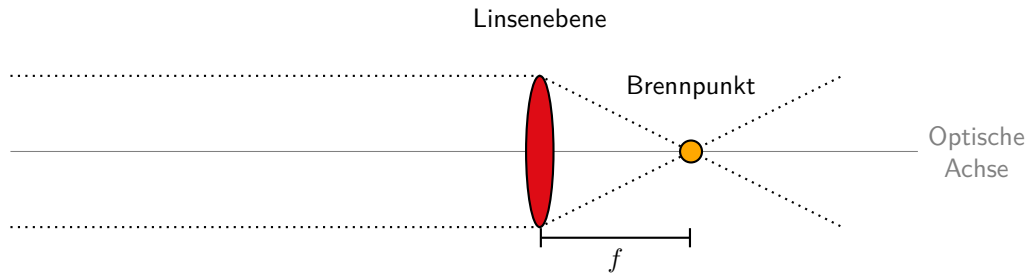


Abbildung 3.4: Brennpunkt

Ist das nicht der Fall, werden die Punkte gestreckt auf der Bildebene dargestellt und das Objekt erscheint unscharf. Für eine bestimmte Brennweite werden immer alle Punkte in einer bestimmten Ebene fokussiert. Diese Ebene nennt man Fokusebene. Je weiter ein Punkt von der Fokusebene entfernt ist, desto unschärfer erscheint er.

### 3.3.2 Biologische Hintergründe

Der Mensch nimmt Licht mit dem Auge wahr. Licht, das auf das Auge trifft, gelangt über die Iris, Linse und Pupille auf die Retina, wo verschiedene Rezeptoren das Licht absorbieren und in elektrische Signale umwandeln, die dann über den Sehnerv ans Gehirn gesendet werden (siehe Abbildung 3.6). Wir werden uns jetzt die einzelnen Bestandteile des Auges im Detail anschauen.

#### Linse

Das menschliche Auge hat funktionell eine sehr große Ähnlichkeit mit der optischen Funktionsweise einer Linse (siehe Abschnitt 3.3.1). Dabei ist die Retina die Bildebene und das betrachtete Objekt die Fokusebene. Die Linse fokussiert dann die eintreffenden Lichtstrahlen auf die Retina. Im Gegensatz zu einer Linse aus Glas ist die Linse im Auge flexibel und kann dadurch ihre Brennweite ändern. Das ist auch wichtig, da man sonst nur Objekte in einer festen Entfernung scharf sehen könnte. Die Fähigkeit zur Anpassung der Brechkraft des Auges nennt man Akkommodation.

Akkommodation

#### Beispiel

Die Akkommodation kann man sehr einfach selbst ausprobieren. Strecken Sie dazu den Arm aus und fokussieren Sie einen Ihrer Finger. Nun erscheinen alle Objekte im Hintergrund verschwommen (unscharf). Fokussieren Sie nun ein Objekt in derselben Richtung, das weiter entfernt ist. Nun erscheint Ihr Finger unscharf.

#### Aufgabe 8: Fokus

Die Brennweite der Linse im menschlichen Auge kann durch Akkommodation dy-

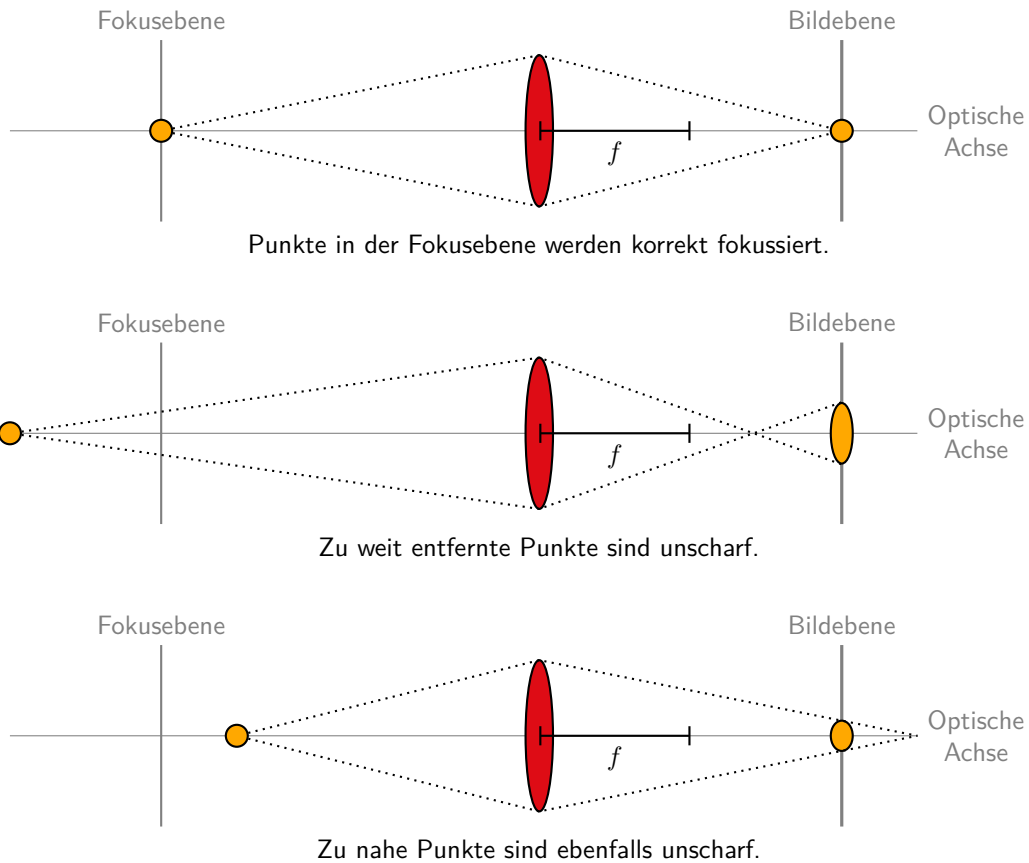


Abbildung 3.5: Fokus

namisch angepasst werden. Wie kommt es, dass trotzdem viele Menschen eine Brille brauchen?

*Tipp:* Manche Brillenträger können mit einiger Anstrengung auch ohne Brille scharf sehen.

## Retina

Die Retina (oder auch Netzhaut) ist die Rückwand des menschlichen Auges, auf der tatsächlich Helligkeit und Farben wahrgenommen werden. Auf ihr werden die eintreffenden Lichtwellen in den Fotorezeptoren in elektrische Signale umgewandelt, die dann über den Sehnerv zum Gehirn gesendet werden. Auf der Retina befinden sich zwei unterschiedliche Arten von solchen Fotorezeptoren:

Fotorezeptoren

### Zapfen *engl. Cones*

Die Zapfen sind für die Farbwahrnehmung zuständig (photopisches Sehen). Es gibt lange, mittlere und kurze Zapfen, die jeweils für andere Wellenlängenbereiche

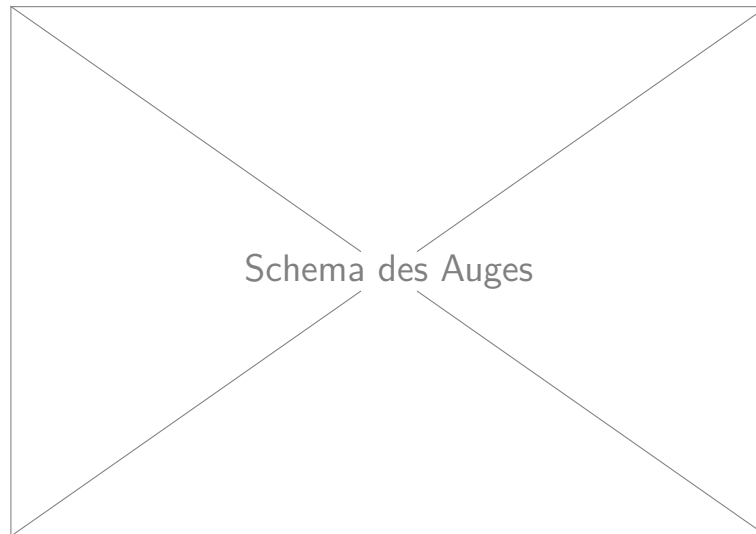


Abbildung 3.6: Schematische Darstellung des menschlichen Auges

empfindlich sind. Diese werden als L-Typ, M-Typ oder S-Typ Zapfen bezeichnet. Man kann die Wellenlängenbereiche der Zapfen ungefähr voneinander abgrenzen:

**S-Typ Zapfen** sind besonders zwischen 400 und 450 nm empfindlich, reagieren aber auf Wellenlängen zwischen weniger als 400 nm und knapp 550 nm.

**M-Typ Zapfen** reagieren auf Licht zwischen 400 nm und 650 nm und sind besonders empfindlich für Wellenlängen um 530 nm.

**L-Typ Zapfen** reagieren auf Licht zwischen 400 nm und 750 nm und sind besonders empfindlich für Wellenlängen um 570 nm.

Es ist auffällig, dass alle Zapfen auf Licht zwischen 470 nm und 540 nm reagieren. Tatsächlich haben Experimente gezeigt, dass Menschen grüne Farben besonders gut sehen können. Der Grund dafür findet sich vermutlich in der Evolution, da es für das Überleben wichtig war, auch kleine Unterschiede in Grüntönen von Pflanzen und eventuell getarnten Feinden zu erkennen. Andere Farben als grün treten in der Natur deutlich seltener auf, sodass eine verstärkte grün-Sensibilität nachvollziehbar ist.

Insgesamt gibt es etwa 7-8 Mio. Zapfen auf der Retina.

### **Stäbchen** *engl. Rods*

Die Stäbchen sind für das Hell-Dunkel-Sehen zuständig (skotopisches Sehen). Stäbchen reagieren deutlich empfindlicher auf Licht, können aber keine Farben unterscheiden. Daher sieht man in sehr dunklen Umgebungen seine Umwelt nur noch in Graustufen. Insgesamt gibt es etwa 100-120 Mio. Stäbchen auf der Retina.

Bevor die Signale von den Stäbchen und Zapfen über den Sehnerv zum Gehirn gelangen, findet noch die retinale Verschaltung statt. Dabei werden die Informationen der

Retinale Verschaltung

unterschiedlichen Zapfentypen sowie der Stäbchen zu einem Gesamteindruck kombiniert, der letztlich das Bewusstsein erreicht. Dies geschieht über die additive Farbmischung (siehe Abschnitt 3.3.5).

Additive  
Farbmischung

Zapfen und Stäbchen sind nicht gleichmäßig auf der Retina verteilt. Abbildung 3.7 zeigt die charakteristische Verteilung von Zapfen (blau) und Stäbchen (rot).

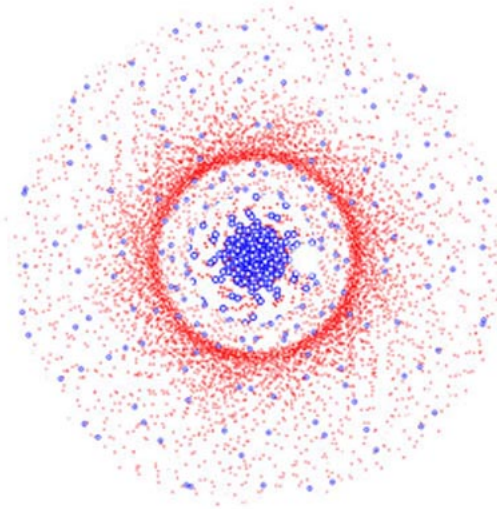


Abbildung 3.7: Zapfen und Stäbchen auf der Netzhaut

In der Abbildung sieht man, dass offensichtlich die meisten Stäbchen (und damit die Farbwahrnehmung) genau in der Mitte der Retina, dem fovealen Bereich stattfindet. Außerhalb dieses Bereiches nimmt die Anzahl der Zapfen stark ab und es befinden sich fast ausschließlich Stäbchen auf der Netzhaut. Die Auswirkung dieser Verteilung ist, dass wir vor allem im direkten Sichtfeld Farben wahrnehmen können. Je weiter ein Objekt von der Sichtachse entfernt ist, desto schlechter können wir dessen Farbe beurteilen. Meistens fällt das jedoch im Alltag nicht auf, da das Gehirn immer ein Bild der Umgebung abspeichert und dieses nach Bedarf aus der Erinnerung mit Farben ergänzt.

Fovea

#### Aufgabe 9: Farbschwächen

Ca. 9 % der Menschheit leiden unter Farbschwächen. Am häufigsten ist die Rot/Grün-Schwäche, bei der die Betroffenen rote und grüne Farben nicht unterscheiden können. Welche Ursachen sind dafür denkbar?

Auf der Netzhaut gibt es außerdem einen kleinen Bereich, in dem sich weder Zapfen noch Stäbchen befinden. Diesen Bereich nennt man den Blinden Fleck. Dort befinden sich keine Rezeptoren, da hier die Nervenbahnen aller Rezeptoren auf der Retina zusammenlaufen und das Auge gebündelt als Sehnerv verlassen. Der Blinde Fleck befindet sich etwa zwischen den größten Ansammlungen von Zapfen und Stäbchen (knapp 20 % von der optischen Achse entfernt). Auch für den Blinden Fleck gilt, dass wir aus dem

Blinden Fleck

Grund kein „Loch“ sehen, da das Gehirn ein temporäres Bild speichert und dadurch solche Löcher aus der Erinnerung füllt. Außerdem befindet sich der Blinde Fleck des linken Auges woanders als der des rechten Auges, sodass das wahrgenommene Bild auch so vervollständigt werden kann.

### 3.3.3 Sehschärfe

#### Definition: Sehschärfe

Die Sehschärfe bezeichnet die Fähigkeit, Konturen und Muster zu erkennen. Sie hängt ab von:

- Auflösungsvermögen
- Abbildungsqualität
- Optische Eigenschaften des Objekts
- Umgebung (Kontrast, Farbe, Form, ...)

Empirische Untersuchungen zur Sehschärfe haben ergeben, dass die Sehschärfe ähnlich wie die Farbwahrnehmung mit zunehmender Entfernung von der optischen Achse schnell abnimmt. Schon bei einer Entfernung von  $5^\circ$  von der optischen Achse beträgt die Sehschärfe nur noch rund 50%. Nur genau in der Mitte des Sichtfeldes ist eine Sehschärfe von 100% möglich. Diesen Bereich nennt man auch Fovea oder auch Gelber Fleck. In Abbildung 3.8 ist die Abnahme der Sehschärfe um die Fovea anschaulich dargestellt.

Fovea  
Fovea

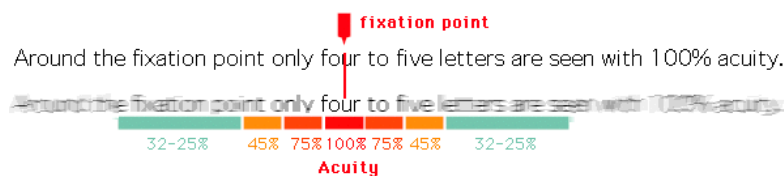


Abbildung 3.8: Foveales Sehen

### 3.3.4 Größenwahrnehmung

Betrachtet man ein Foto von zwei Menschen, kann man meistens relativ leicht feststellen, welche der beiden Personen größer ist. Das funktioniert auch, wenn eine Person näher an der Kamera steht als die andere. Das legt nahe, dass die wahrgenommene Größe nicht alleine von der tatsächlichen Größe der Person auf dem Foto abhängt. Die tatsächliche Größenwahrnehmung wird durch das Emmertsche Gesetz beschrieben:

Emmertsches Gesetz

#### Definition: Emmertsches Gesetz

Die Wahrgenommene Größe  $H$  eines Objektes hängt von der Größe des Objektes



auf der Retina  $B$ , der wahrgenommenen Distanz  $D$  zum Objekt und der Entfernung zwischen Linse und Retina  $W$  ab. Es gilt:

$$\frac{B}{H} = \frac{W}{D} \Rightarrow H \sim B \cdot D$$

Wie man sieht, arbeitet das Gesetz mit der wahrgenommenen Größe und Distanz. Es ist wichtig, sich klar zu machen, dass man eine Distanz oder Größe nicht absolut wahrnehmen kann, sondern nur abschätzt (wie die wahrgenommene Distanz ermittelt wird, wird im Abschnitt 3.3.7 erklärt.). In den meisten Fällen kann man außerdem  $W$  als konstant ansehen (ungefähr 2,54 cm). Am besten kann man sich die Wirkung des Emmertschen Gesetzes durch visuelle Illusionen veranschaulichen:

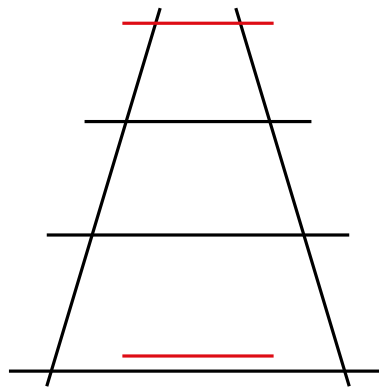


Abbildung 3.9: Ponzo Illusion

In der Ponzo Illusion (Abbildung 3.9) sind die beiden roten Linien genau gleich lang. Trotzdem scheint die untere deutlich kürzer zu sein. Dies kann man mit dem Emmertschen Gesetz begründen. Die retinale Größe  $B$  der beiden Linien ist genau gleich groß. Allerdings erwecken die „Schienen“ den Eindruck, als ob die obere Linie weiter weg ist. Damit erhöht sich die wahrgenommene Distanz  $D$ . Da  $H \sim B \cdot D$  ist, erhöht sich so auch die wahrgenommene Größe.

Ponzo Illusion

#### Aufgabe 10: Mond-Illusion

Erklären Sie mithilfe des Emmertschen Gesetzes, wieso der Mond am Horizont oft größer erscheint, als direkt über einem?

### 3.3.5 Farbmodelle

Thomas Young und Hermann von Helmholtz haben aus der Tatsache, dass der Mensch nur drei unterschiedliche Zapfentypen besitzt, geschlossen, dass sich umgekehrt auch jede wahrnehmbare Farbe als Kombination von drei verschiedenen Wellenlängen herstellen lassen muss. Es hat sich herausgestellt, dass nicht nur diese Theorie stimmt, sondern der

Mensch tatsächlich nicht unterscheiden kann, ob eine bestimmte Farbe durch Überlagerung anderer Wellenlängen oder als „reine“ Wellenlänge entsteht. Diese Überlagerung von Farben bezeichnet man als Farbaddition. Die physikalische Grundlage ist die Additive Farbmischung<sup>2</sup>.

Farbaddition  
Additive  
Farbmischung

Für die Farbaddition gilt folgendes:

1. Falls weder Stäbchen noch Zapfen durch Licht erregt werden, sehen wir schwarz.
2. Werden alle Stäbchen und Zapfen in etwa gleichem Maße erregt, sehen wir weiß.
3. Die Farbe, die durch die Überlagerung von zwei (oder mehr) verschiedenen Wellenlängen entsteht, ist nicht von einer „reinen“ Farbe zu unterscheiden, die dieselben Stäbchen und Zapfen erregt.

Diese Erkenntnisse finden in der Informatik exzessiv Anwendung. Von den drei Primärfarben, die vom Menschen wahrgenommen werden lässt sich nun das RGB-Farbmodell ableiten. Dieses stellt Farben als eine Kombination eines Rot-, Grün- und Blau-Anteils dar, die dann additiv gemischt werden<sup>3</sup>. Formal ist eine Farbe im RGB-Farbmodell also ein 3-Tupel, wobei die erste Komponente der Rotanteil, die zweite der Grünanteil und die dritte der Blauanteil ist. Üblicherweise verwendet man für jede Farbkomponente 1 Byte zum Speichern, sodass also Zahlen zwischen 0 und 255 möglich sind. Damit entspricht die Farbe (255, 0, 0) reinem rot und (100, 255, 0) einer Mischung aus Rot und grün.

RGB-Farbmodell

Das RGB-Farbmodell wird auch bei LCD-Displays angewendet. In LCD-Displays besteht jeder einzelne Pixel aus drei Subpixeln (je einer pro Farbkanal). Diese Subpixel sind so klein, dass man beim Betrachten des Bildschirms die einzelnen Farben nicht unterscheiden kann, sondern nur die Farbe erkennt, die durch die Überlagerung der Farben der Subpixel entsteht. Andere Monitorarten (z.B. Plasma- oder OLED-Displays) verwenden zwar unterschiedliche Technologien, arbeiten aber auch mit dem RGB-Farbmodell.

Subpixel

#### ✓ Übung: RGB-Farbmodell

Öffnen Sie ein Grafikprogramm und suchen Sie, wo Sie die Farbe für Ihre Zeichnung auswählen können. Experimentieren Sie mit den Farbreglern im RGB-Farbmodell. Versuchen Sie die Primärfarben rot, grün und blau, aber auch andere Farben wie lila, orange oder magenta einzustellen. Wie erhalten Sie eine dunklere Version einer Farbe?

Versuchen Sie ohne weitere Hilfsmittel (also durch Ausprobieren) die Farbe [dieses Textes](#) einzustellen.

Was bei der obigen Übung aufgefallen sein sollte, ist, dass das RGB-Farbmodell nicht besonders intuitiv ist, obwohl es auf der Funktionsweise unserer Augen basiert. Im groben gelingt es uns zwar, eine Farbe abzuschätzen, aber im Detail ist dies für uns sehr

<sup>2</sup>Es gibt auch eine Subtraktive Farbmischung. Diese findet im Tuschkasten statt. Mischt man dort alle Farben, erhält man braun bzw. schwarz. Weiß erhält man, indem man keine Farbe aufträgt. Die Subtraktive Farbmischung ist für uns aber nicht weiter interessant.

<sup>3</sup>Die Primärfarben bei der Additiven Farbmischung sind andere als bei der Subtraktiven Farbmischung. Daher sind die Primärfarben nicht Rot, Gelb und Blau, sondern Rot, Grün und Blau.

schwer. Auch die Tatsache, dass Helligkeit in dem Farbmodell nicht direkt repräsentiert ist, widerspricht unserer Intuition. Um dieses Problem zu lösen, gibt es einige andere Farbmodelle. Ein Beispiel ist das HSV-Farbmodell (Hue-Saturation-Value). Auch hier wird eine Farbe durch ein 3-Tupel repräsentiert. Allerdings handelt es sich hier nicht mehr um Rot-, Grün- und Blauanteil, sondern um die Farbe (Hue), die Sättigung (Saturation) und die Helligkeit (Brightness oder Value). Der Hue-Wert wird als Winkel in einem Farbrad interpretiert.

HSV-Farbmodell

#### ✓ Übung: HSV-Farbmodell

Experimentieren Sie wie schon beim RGB-Farbmodell mit den Einstellungen für das HSV-Farbmodell. Wie können Sie hier eine Farbe aufhellen oder abdunkeln? Wie nahe kommen Sie an die Farbe [dieses Textes](#) heran?

### 3.3.6 Helligkeits- und Kontrastwahrnehmung

Um zu sehen, wie die Wahrnehmung von Helligkeit und Kontrast funktioniert, müssen wir zunächst einige Begriffe klären:

#### Definition: Helligkeit

- Der Betrag der Lichtenergie, der eine Lichtquelle verlässt, ist dessen **Strahlung** (engl. *Radiance*).
- Der Betrag der Lichtenergie, die ein Objekt erreicht, bestimmt dessen **Beleuchtung** (engl. *Illuminance*).
- Das vom Objekt reflektierte Licht bestimmt die **Lichtstärke/Luminanz** (engl. *Luminance*).
- Der Eindruck, der am Ende im Bewusstsein entsteht ist die **Helligkeit**.

Die wahrgenommene Helligkeit wird durch die Stimulation der Stäbchen auf der Retina hervorgerufen und ist von der Luminanz abhängig.

Stäbchen

Die Helligkeit ist also ein subjektiver Eindruck. Als physikalische Größe wird sie auch nicht direkt gemessen. Stattdessen misst man die Luminanz. Dass Helligkeit keine objektiv messbare Größe ist, betrachten wir weiter unten.

Eine sehr eng mit der Helligkeit verwandte Größe ist der Kontrast. Er lässt sich folgendermaßen definieren:

#### Definition: Kontrast

Der Kontrast ist das Verhältnis der Luminanz eines Objektes ( $L_V$ ) zu dessen Hin-

tergrun ( $L_H$ ). Er berechnet sich folgendermaßen:

$$K = \frac{L_V - L_H}{L_H} = \frac{\Delta L}{L_H} \quad \Delta L = L_V - L_H$$

Die Kontrastsensitivität des Menschen liegt bei etwa 1:200 bis 1:250. Das heißt, dass bis zu 250 verschiedene Graustufen unterschieden werden können. Diese Verhältnisse sind überraschend gering. Tatsächlich liegt der Gesamtkontrast des Auges bei etwa 1:100.000. Dieses Kontrastverhältnis kann aber nicht gleichzeitig erkannt werden. Stattdessen reguliert die Iris die Menge des einfallenden Lichts, sodass zu einem Zeitpunkt immer nur bis zu 250 Graustufen unterschieden werden können. Das folgende Beispiel veranschaulicht die Funktionsweise der Kontrastwahrnehmung:

Kontrastsensitivität

Gesamtkontrast

### Beispiel

Geht man in der Nacht (bzw. in der Dämmerung) auf einem beleuchteten Waldweg, scheint der Wald neben dem Weg eine einzige schwarze Wand zu sein. Man kann wenige Meter entfernt weder Blätter noch Äste oder Baumstämme erkennen. Ist die Beleuchtung aber abgeschaltet, kann man plötzlich die Umrisse weit entfernter Bäume erkennen.

Der Grund dafür ist, dass sich das Auge an die Dunkelheit angepasst hat. In beiden Fällen, sieht man nur etwa 250 verschiedene Graustufen. Ist Beleuchtung aus, kann man also die Bäume vom leicht dunkleren Hintergrund unterscheiden. Ist aber die Beleuchtung an, passt sich das Auge entsprechend an. Man kann weiterhin 250 Graustufen unterscheiden, allerdings ist die Abstufung der jetzt viel weniger fein. Die Details des dunklen Waldes gehen dabei in der Perzeption verloren.

Wie gut man Kontraste wahrnehmen kann, hängt neben der Umgebungshelligkeit aber auch noch von weiteren Faktoren ab:

**Kontrastintensität** Der Mensch kann vor allem starke Kontraste sehr gut erkennen.

Je geringer die Helligkeitsunterschiede sind (je geringer der Kontrast ist) desto schwerer fällt es, überhaupt Unterschiede zu erkennen.

**Räumliche Frequenz** Die räumliche Frequenz beschreibt die Häufigkeit des Wechsels zwischen Hell und Dunkel auf einer Fläche. Allgemein ist die Kontrastwahrnehmung bei einem Augenabstand von ca 0,5m bei einer Räumlichen Frequenz von  $1 \text{ cm}^{-1}$  bis  $10 \text{ cm}^{-1}$  am besten. Hier können auch noch sehr geringe Lichtintensitäten unterschieden werden.

In der Mensch-Computer-Interaktion werden diese Erkenntnisse vor allem Im Bereich der Bildverarbeitung sehr viel angewendet. So kann man deutlich effizientere Bildkompressionsalgorithmen entwickeln, indem man berücksichtigt, dass Graustufen ab einem bestimmten Schwellenwert gar nicht unterschieden werden können<sup>4</sup>.

<sup>4</sup>Die JPEG-Komprimierung nutzt sehr ähnliche Erkenntnisse aus, um Bilder auch bei hoher Qualität sehr gut zu komprimieren.

**Machsche Bänder** Machsche Bänder sind ein visueller Wahrnehmungseffekt, der zwischen Kanten entsteht, die leicht unterschiedlich schattiert sind. Genau an diesen Kanten scheint der Kontrast stark erhöht zu sein. Dieser Effekt veranschaulicht sehr gut die Fähigkeit des menschlichen Auges, Unterschiede wahrzunehmen. Auch wenn es schwer fällt, durch reine Betrachtung mit dem Auge quantitative Angaben über einen Farbverlauf zu machen, ist es relativ einfach zu erkennen, an welcher Stelle der Verlauf beginnt bzw. endet.

Der biologische Hintergrund zu den Machschen Bändern ist die sogenannte Laterale Hemmung. Die Idee dabei ist, dass die einzelnen Fotorezeptoren nicht unabhängig voneinander arbeiten, sondern direkt mit ihren Nachbarn verschaltet sind. Je stärker ein Rezeptor stimuliert wird, desto stärker ist auch die Hemmung seiner Nachbarn. Die Auswirkung dieses Modells sind genau die Machschen Bänder: An den Stellen, an denen benachbarte Rezeptoren sehr unterschiedlich stimuliert werden, entsteht eine ungewöhnlich hohe bzw. niedrige Stimulation der Nervenzellen.

Laterale Hemmung

### 3.3.7 Objekterkennung

Im Folgenden beschäftigen wir uns mit dem Prozess der Objekterkennung beim Menschen. Wie schon beim Stufenmodell (Abbildung 3.1) gibt es auch hier sowohl Bottom-Up- als auch Top-Down-Processing. Hier betrachten wir vor allem das Bottom-Up-Processing. Die Top-Down-Variante funktioniert dabei analog zum Stufenmodell der Wahrnehmung: Vorhergegangene Informationen beeinflussen also die Interpretation von neu wahrgenommenen Strukturen.

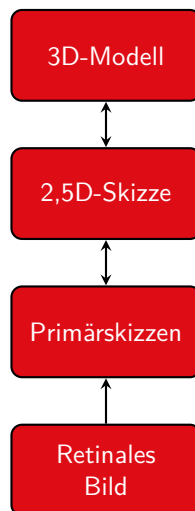


Abbildung 3.10: Objekterkennung: Stufenmodell

In der Forschung haben sich für die Objekterkennung zwei gegensätzliche Erkennungsmodelle herausgebildet:

**Ansichtsbasierte Erkennungsmodelle** Diese Erkennungsmodelle gehen davon aus, dass

das Gehirn die wahrgenommenen Bildern mit „Schnappschüssen“ bekannter Objekte vergleicht und so feststellt, welches Objekt bzw. welche Objekte man gerade betrachtet. Hier geht man davon aus, dass im Gehirn grundsätzlich bildbasierte Repräsentationen von Objekten (als 2D-Repräsentationen) gespeichert sind.

**Strukturelle Erkennungsmodelle** Diese Erkennungsmodelle gehen davon aus, dass das Gehirn jedes Objekt zuerst in seine grundlegenden Strukturen zerlegt und diese dann (als 3D-Repräsentation) wiedererkennt.

Abbildung 3.11 sind diese beiden Ansätze veranschaulicht. Wir beschäftigen uns hier mit dem Ansatz der strukturellen Erkennungsmodelle.

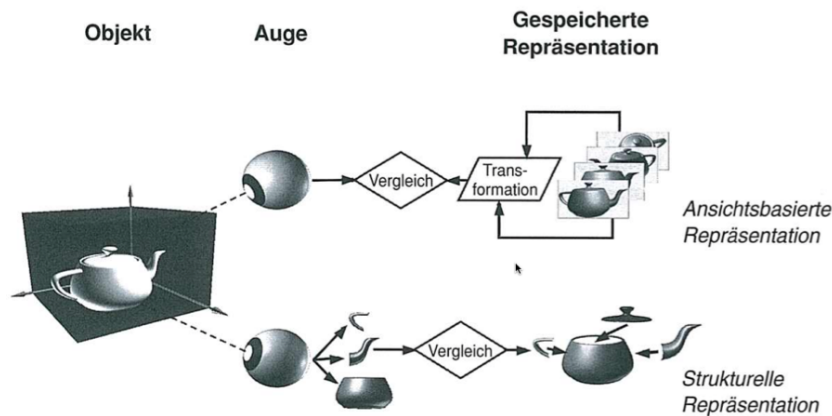


Abbildung 3.11: Objekterkennungsmodelle

**Retinales Bild** Das Retinale Bild sind die *Rohdaten*, die von den Rezeptoren im Gehirn ankommen. Vorstellen kann man sich dies wie bei einem Computerbildschirm. Das Retinale Bild sind die Daten für die einzelnen Pixel.

**Primärskizze** Aus dem Retinalen Bild erstellt das Gehirn dann zwei Primärskizzen:

**Rohe Primärskizze engl. Raw Primal Sketch** Auf dem Retinalen Bild werden zunächst Kanten unterschiedlicher Helligkeit und Farbe, Flecken und Linien extrahiert. Diese Linien sind zunächst noch ungeordnet und können abhängig von den Lichtverhältnissen und der Qualität des Retinalen Bildes (Entfernung zum Objekt, Sichtverhältnisse usw.) Lücken und Fehler enthalten.

**(Volle) Primärskizze engl. Full Primal Sketch** In der rohen Primärskizze werden dann Fehler korrigiert und Lücken gefüllt. Zum Ist ein Objekt zum Beispiel zum Teil verdeckt, dann kann das Gehirn in diesem Schritt raten, wie die Kante weiter verläuft. Am Ende ist so eine Primärskizze vorhanden, die die zweidimensional wahrgenommenen Konturen des Objektes Beschreibt.

**2,5D-Skizze** Bei der Erstellung der 2,5D-Skizze wird das zweidimensionale Bild der Primärskizze mit Tiefeninformationen verbunden. So soll eine Repräsentation entstehen, die aus orientierten Oberflächen mit Richtungszuweisungen zusammengesetzt ist und vom Blickwinkel des Betrachters unabhängig ist. Die 2,5D-Skizze enthält allerdings nur Informationen, die aus der Primärskizze hervorgehen. Das bedeutet insbesondere, dass die Rückseite des Objektes noch unbekannt bleibt. Das Erstellen der 2,5D-Skizze besteht aus zwei Schritten:

**Flächenerkennung** Durch Tiefenhinweise (z.B. Textur und Schattierung) schätzt das Gehirn die Größe und Orientierung von Flächen ab. Es gibt sehr viele verschiedene Tiefenhinweise (s. Tabelle 3.1), sodass normalerweise mehrere Hinweise kombiniert werden, um eine möglichst realistische Einschätzung zu liefern. Tiefenhinweise werden sowohl für die Einschätzung absoluter und relativer Größe als auch für das Abschätzen von Entfernungen verwendet (vgl. Emmertsches Gesetz).

Tiefenhinweise

**Kanteninterpretation** An den Kanten, die Oberflächen miteinander verbinden, stellt das Gehirn fest, wie die Flächen ausgerichtet sind. Es unterscheidet konvexe und konkave Orientierungskanten sowie Tiefenkanten. Tiefenkanten sind Kanten, die dadurch entstehen, dass ein Objekt ein anderes Verdeckt.

Die Erstellung der 2,5D-Skizze ist der Schritt der Objekterkennung, bei dem dem Gehirn die meisten Fehler unterlaufen. Das liegt einerseits daran, dass es physikalisch-biologisch einfach nicht möglich ist, ein Objekt auf Anhieb korrekt zu erkennen und andererseits daran, dass das Gehirn sehr viele Heuristiken anwenden muss, um die Oberflächen und Kanten mithilfe der Tiefenhinweise richtig zu erkennen. Sind die Lichtverhältnisse zum Beispiel gerade ungünstig, so kann es schnell passieren, dass die 2,5D-Skizze gravierende Fehler enthält. Das Gehirn versucht auf Grundlage der vorhergehenden Lebenserfahrung, eine möglichst plausible 2,5D-Repräsentation zu erzeugen.

**3D-Modell** Im letzten Schritt wird aus der 2,5D-Skizze ein 3D-Modell erstellt. Dazu müssen also fehlende Kanten und Flächen sowie die Rückseite des Objektes ergänzt werden. Strukturelle Erkennungsmodelle gehen davon aus, dass dafür die 2,5D-Skizze in verschiedene Einzelstrukturen zerlegt wird, die dann unabhängig voneinander ergänzt und danach wieder zusammengesetzt werden. Wie diese Einzelstrukturen genau aussehen ist nicht klar. Allerdings gibt es verschiedene Theorien:

**Generalisierte Zylinder** D. Marr stellte 1982 die Theorie vor, dass das Gehirn jedes Objekt in verschiedenartig geformte Zylinder zerlegt. Ein Mensch könnte zum Beispiel aus 6 Zylindern bestehen: Kopf, Korpus, 2 Arme und 2 Beine. Im Detail besteht dann jeder Arm wiederum aus Ober- und Unterarm (2 Zylinder) und die Hand aus der Handfläche sowie den Fingern.

**Geonen-Theorie** Diese Theorie basiert auf Marrs Ansatz, erlaubt aber neben Zylindern auch noch andere geometrische Primitive. Diese Primitive werden als Geone bezeichnet. Geone zeichnen sich vor allem durch Projektionsinvarianz aus (d.h. sie

Geone

Tiefenhinweis	Wirkungsbereich	Klassifizierung	Positionsbestimmung
Verdeckung	Kompletter Bereich	Monokular	Relativ
Disparität	Bis 10 m	Binokular	Relativ
Konvergenz	Bis 2 m	Binokular	Absolut
Akkommodation	Bis 2 m	Monokular	Absolut
Image Blur	Kompletter Bereich	Monokular	Relativ
Linearperspektive	Kompletter Bereich	Monokular	Absolut
Texturgradient	Kompletter Bereich	Monokular	Relativ
Relative Größe	Kompletter Bereich	Monokular	Absolut
Bekannte Größe	Kompletter Bereich	Monokular	Absolut
Höhe im Gesichtsfeld	Über 30 m	Monokular	Relativ
Atmosphärische Perspektive	Über 30 m	Monokular	Relativ
Shape from Shading	Kompletter Bereich	Monokular	Relativ
Schattenwurf	Kompletter Bereich	Monokular	Relativ
Bewegungsparallaxe	Über 20 m	Dynamisch	Relativ
Accretion	Kompletter Bereich	Dynamisch	Relativ

Tabelle 3.1: Tiefenhinweise

sind aus unterschiedlichen Blickwinkeln nicht verzerrt). Ursprünglich hat I. Biederman etwa 36 verschiedene Geone definiert.

### 3.3.8 Gesichtserkennung

Gesichtserkennung hat sich evolutionär als essentieller Fähigkeit zum Überleben herausgestellt. Im Kampf ums Überleben ist es extrem von Vorteil, wenn man mögliche Feinde sofort an deren Gesicht erkennen kann. Daher unterscheidet sich die Gesichtserkennung auch fundamental von der Objekterkennung.

Im Gehirn gibt es eine besondere Spindelwindung (Gyrus fusiformis) in der Großhirnrinde, von der vermutet wird, dass sie allein zur Gesichtserkennung da ist. Vor allem führt der Defekt dieser Windung zur Gesichtsbblindheit, einer Krankheit, bei der die Betroffenen die Fähigkeit verlieren, Gesichter zu erkennen. Da dieser Gehirnteil offenbar allein für das Erkennen von Gesichtern zuständig ist, erkennt das Gehirn Gesichter sehr schnell.

Gyrus fusiformis



## 3.4 Auditive Wahrnehmung

Die Ohren sind die Sekundären Wahrnehmungsorgane des Menschen. Das Hören hat sich evolutionär besonders als Werkzeug zur Kommunikation mithilfe von Sprache herausgebildet.

### 3.4.1 Physikalische Hintergründe

#### Definition: Schallwellen

Schallwellen sind Longitudinalwellen, die sich im Prinzip in jedem Medium ausbreiten können. Hier beschäftigen wir uns nur mit Schallwellen, die sich in der Luft ausbreiten.

Schallwellen breiten sich um einen Entstehungspunkt konzentrisch im Raum aus. Als Longitudinalwellen schwingen die einzelnen Oszillatoren (in diesem Fall die Luftmoleküle) in Richtung der Ausbreitungsrichtung, wodurch Luftdruckschwankungen entstehen. Die Ausbreitungsgeschwindigkeit von Schall unterscheidet sich je nach Medium sehr stark, ist aber deutlich geringer als die von Licht. In Luft breitet sich Schall mit 331 m/s bei 0 °C und mit 343 m/s bei 20 °C aus.

#### Aufgabe 11: Schallgeschwindigkeit

Zwei Personen stehen 500 m voneinander entfernt. Person A klatscht in die Hände. Wie lange dauert es, bis das Klatschgeräusch Person B erreicht? Wie lange dauert es, bis Person B das Zusammentreffen der Hände von Person A sieht?

*Rechnen Sie mit der Schallgeschwindigkeit bei 20 °C.*

Als Welle haben auch Schallwellen eine Amplitude und eine Frequenz:

**Die Frequenz** definiert die Tonhöhe. Je schneller die einzelnen Luftmoleküle schwingen, desto höhere Töne hört man. Das gesamte Frequenzspektrum der hörbaren Töne geht von etwa 20 Hz bis 20 kHz (entsprechend Wellenlängen von ca. 1,7 cm bis 17 m). Darüber befindet sich der Ultraschallbereich.

**Die Amplitude** ist die Lautstärke. Eine hohe Amplitude entspricht auch einer hohen Lautstärke. Die Lautstärke entspricht der Stärke der Luftdruckänderung und wird in Dezibel (dB) gemessen. Dezibel (1 dB = 0,1 B) ist eine logarithmische Maßeinheit zur Kennzeichnung von Pegeln. Sie ist definiert über das Verhältnis der Energiegrößen  $P_2$  und  $P_1$ .  $P_1$  ist dabei der Schalldruck bei der Hörschwelle, also etwa  $2 \cdot 10^{-5}$  Pa.

$$L = \log_{10} \left( \frac{P_2^2}{P_1^2} \right) \text{ B} = 20 \cdot \log_{10} \left( \frac{P_2}{P_1} \right) \text{ dB}$$

Eine Konsequenz dieses Logarithmischen Verhältnisses ist, dass bereits etwa alle 10 dB eine Verdoppelung der Lautstärke wahrgenommen wird.

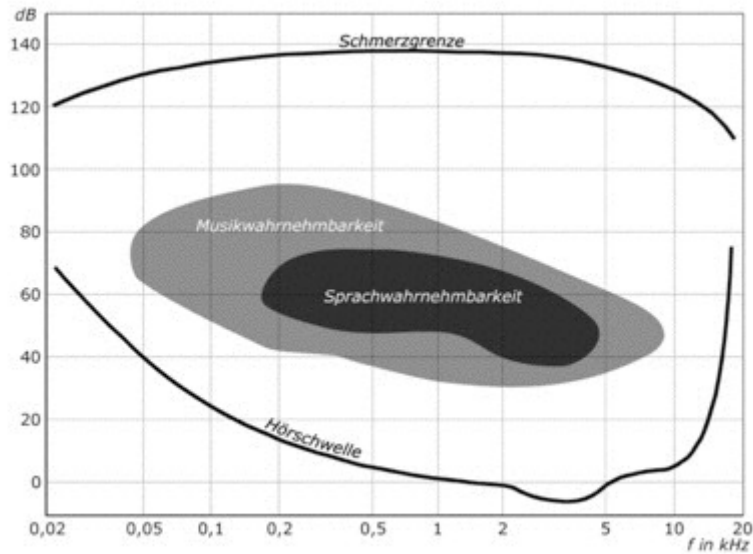


Abbildung 3.12: Auditive Hörschwelle und Schmerzgrenze

Genauso wie beim Licht gibt es Grenzen der Wahrnehmung von Schall. Analog zur Farb- und Helligkeitswahrnehmung, können wir bei einer bestimmten Lautstärke besonders viele Frequenzen unterscheiden. Im Detail ist das in Abbildung 3.12 dargestellt. Dort ist auch der Frequenz- und Lautstärkebereich von Musik und Sprache dargestellt, in dem der Mensch besonders sensibel ist.

**Grundtöne und Obertöne** Im Gegensatz zum Licht unterscheiden wir bei akustischen Reizen zwischen unregelmäßigen und periodischen Luftdruckschwankungen. Nur periodische Luftdruckschwankungen werden sauber als Ton wahrgenommen. Unregelmäßige Schwankungen werden nur als Geräusche wahrgenommen. Töne werden dadurch erzeugt, dass ein Gegenstand (z.B. eine Saite oder die Stimmbänder) in Schwingung versetzt wird. Diese Schwingungen mit einer bestimmten Grundfrequenz (Grundton) werden dann reflektiert, wodurch Schwingungen der n-fachen Grundfrequenz entstehen. Diese nennt man dann n-te Harmonische oder (n-1)-ter Oberton.

Grundton

Oberton

**Klangfarbe** Je nach Materialien und Art der Tonerzeugung vermischen sich Grund- und Obertöne auf charakteristische Weise. Das so entstandene Gemisch bezeichnet man als Klangfarbe. Im Computer entspricht die Klangfarbe dem Frequenzspektrum.

Klangfarbe

### 3.4.2 Biologische Hintergründe

Beim Menschen findet die auditive Wahrnehmung (Hören) im Ohr statt. Dort wird die Luftdruckänderung in elektrische Signale umgewandelt, die dann vom Gehirn weiterverarbeitet werden. Das Ohr ist unterteilt in drei Bereiche:

1. **Das Außenohr** beginnt bei der Ohrmuschel und geht über den Gehörgang zum Trommelfell.
2. **Das Mittelohr** beginnt beim Trommelfell und geht über die Gehörknöchelchenkette bis zur Schnecke.
3. **Das Innenohr** besteht aus der Schnecke (Cochlea) und dem Hörnerv.

Cochlea

**Trommelfell** Einer der wichtigsten Bestandteile des Ohres ist das Trommelfell. Dies ist eine etwa 0,1 mm dünne kegelförmige Hautschicht zwischen dem Gehörgang und dem Mittelohr. Luftdruckänderungen führen nun dazu, dass das Trommelfell anfängt zu vibrieren. Diese Vibrationen werden dann über die Gehörknöchelkette zur Schnecke weitergegeben.

**Schnecke** In der Cochlea befinden sich die Rezeptoren, die die Bewegungen des Trommelfells in elektrische Signale umwandeln. Diesen Vorgang nennt man Transduktion. In der Schnecke befindet sich eine Flüssigkeit, die durch die Gehörknöchelkette in Bewegung gesetzt wird. Das sorgt dafür, dass die Rezeptoren (Haarzellen) stimuliert werden. Je nachdem wie hoch der Ton ist, der für den Stimulus gesorgt hat, werden die Haarzellen weiter vorne bzw. weiter hinten in der Schnecke stimuliert. Hohe Töne stimulieren dabei die Haarzellen ganz vorne in der Schnecke, tiefe Frequenzen stimulieren die Haarzellen am Ende der Schnecke.

Cochlea

Transduktion

Haarzellen

### 3.4.3 Geräuschlokalisierung

#### Definition: Geräuschlokalisierung

Geräuschlokalisierung ist die Bestimmung von Richtung und Distanz einer Geräuschquelle relativ zur eigenen Position.

Bei der Geräuschlokalisierung verwendet das Gehirn verschiedene Distanzhinweise (s. Tabelle 3.2).

Distanzhinweise

Die Distanzhinweise funktionieren ähnlich wie die Tiefenhinweise bei der visuellen Wahrnehmung. Allerdings ist die Lokalisierungsfähigkeit des auditiven Systems sehr limitiert. Im Durchschnitt können Menschen horizontal nur ca. 5 Geräuschquellen sinnvoll unterscheiden. Vertikal und in der Tiefe jeweils nur etwa 3 verschiedene Geräuschquellen.

Tiefenhinweise

### 3.4.4 Akustik in der MCI

Für die Mensch-Computer-Interaktion sind die Erkenntnisse über die Funktionsweisen des Gehörs sehr relevant. Analog zu den Bildkomprimierungsalgorithmen, die auf Grundlage der Erkenntnisse über die visuelle Wahrnehmung entwickelt werden konnten, gibt es auch Kompressionsalgorithmen für Audiodaten. Ein bekanntes Beispiel ist das MP3-Format. Hier werden beispielsweise Daten, die außerhalb der menschlichen Hörschwelle

Distanzhinweis	Erläuterung
Interaurale Unterschiede	Phasenverschiebung, Zeitdifferenz, Lautstärkeunterschiede und unterschiedliche Dämpfung der beiden Ohren
Geräuschspektrum	Entfernte Geräusche klingen gedämpfter
Lautstärke	Entfernte Geräusche sind leiser
Bewegungsparallaxe	Nahe Geräusche bewegen sich schneller am Hörer vorbei als entfernte
Reflektionen	In geschlossenen Räumen erreicht ein direktes Geräusch den Hörer schneller als die Reflektion. Das Verhältnis gibt Aufschluss über die Distanz.

Tabelle 3.2: Distanzhinweise

liegen, gar nicht abgespeichert. Ein anderer Effekt, der hier ausgenutzt wird, ist die Maskierung von Tönen. Diese tritt auf, wenn ein Klang durch ein deutlich lauterer Geräusch überdeckt wird. In einem gewissen Frequenzbereich um das laute Geräusch kann das menschliche Gehör keine anderen Geräusche wahrnehmen, sodass diese im MP3-Format ebenfalls ausgefiltert werden. Maskierung funktioniert besonders bei hohen Tönen sehr gut.

**Auditory Icon** Da die auditive Wahrnehmung der sekundäre Weg der Wahrnehmung des Menschen ist, eignet er sich besonders gut, um unterstützende Signale zu senden. Dies nennt man Auditory Icon. So ein Icon ist ein akustisches Symbol bzw. Signal, welches ein Objekt oder einen Vorgang auf dem User Interface darstellt. Meistens werden dafür natürliche Geräusche wie zum Beispiel raschelndes Papier (Papierkorb) oder das Läuten einer Glocke verwendet.

Auditory Icon

**Earcon** Ein Earcon ist eng verwandt mit Auditory Icons. Ein Earcon ist allerdings ein synthetisches Geräusch zur abstrakten Codierung von Informationen im User Interface. Da es sich hier nicht um ein natürliches Geräusch handelt, muss dessen Bedeutung zuerst erlernt werden. Ein Beispiel für ein Earcon ist das Abstands-Piepen eines Autos beim Einparken.

Earcon

### 3.4.5 Vestibulärer Sinn

Der vestibuläre Sinn (umgangssprachlich Gleichgewichtssinn) befindet sich ebenfalls im Ohr. Er nutzt die Eigenschaften der Flüssigkeit in der Cochlea aus, um lineare Beschleunigung, Drehbeschleunigung und die Richtung der Erdanziehungskraft festzustellen. Dieser Sinn spielt auch für die visuelle Wahrnehmung eine wichtige Rolle. Durch die Kombination des visuellen und vestibulären Sinnes, kann das Gehirn das Bild auf der Retina sehr gut stabilisieren, während sich der Kopf bewegt. Diese Stabilisierung nennt man Vestibulookulären Reflex.

Cochlea

Vestibulookulärer Reflex

Der vestibuläre Sinn steht oft im Widerspruch zu anderen Sinnen, insbesondere dem visuellen Sinn. Das Gehirn schafft es meistens, die beiden Sinne so zu kombinieren, dass ein sinnvolles Resultat entsteht. Ein Beispiel dafür sind Computerspiele. Auch wenn die visuelle Wahrnehmung erkennt, dass man sich bewegt, ist durch den vestibulären Sinn klar, dass sich nur die Spielfigur bewegt und nicht man selbst. Allerdings klappt das nicht immer. Wenn am Bahnhof der Zug auf dem Nachbargleis losfährt, dann passiert es leicht, dass man für einen Moment denkt, dass es der eigene Zug ist, der sich bewegt.

## 3.5 Tastsinn

### Definition: Tastsinn

Der Tastsinn umfasst diverse Wahrnehmungsaspekte:

- Druckwahrnehmung
- Temperaturwahrnehmung
- Schmerzwahrnehmung
- Viszeroseption
- ...

Im Kontext der MCI werden wir den Tastsinn nur kurz behandeln. Das liegt hauptsächlich daran, dass es bisher nicht viele realistischen Ein- bzw. Ausgabegeräte gibt, die den Tastsinn verwenden.

Der Tastsinn wird unterteilt in taktile und haptische Wahrnehmung:

**Taktile Wahrnehmung** beschreibt passives Berührtwerden. Sie basiert auf unter der Haut liegenden Rezeptoren, die auf Stimuli wie Druck (Mechanorezeptoren), Wärme (Thermorezeptoren) oder Schmerz (Nozizeptoren) reagieren.

**Haptische Wahrnehmung** beschreibt aktives Erkunden durch den Tastsinn. Dabei werden dieselben Rezeptoren wie bei der taktilen Wahrnehmung verwendet. Darüber hinaus gibt es noch weitere kinästhetische Rezeptoren der Muskeln, Sehnen und Gelenken, die verwendet werden.

Auch beim Tastsinn haben die Rezeptoren bestimmte Bereiche, in denen die Wahrnehmung möglich ist. Vibrationen können zum Beispiel im Bereich von 3 bis 500 Hz wahrgenommen.

### 3.5.1 Propriozeption

Propriozeption ist die Wahrnehmung der eigenen Körperbewegung (Kinästhesie) und -lage (Lagesinn) im Raum. Die Propriozeption ist für die MCI sehr wichtig, da es ohne die-

sen Sinn für den Menschen nicht möglich wäre, eine Maus zu benutzen, ohne gleichzeitig die eigene Bewegung visuell zu verfolgen.

#### **Beispiel: Haptische Geräte**

Da die Interaktion zwischen Mensch und Computer zu großen Teilen durch die Hände des Menschen gesteuert wird, kann man Tastatur und Maus zu haptischen Eingabegeräten zählen. Diese nutzen allerdings die Möglichkeiten des Tastsinns nur begrenzt aus. Es aber auch andere Geräte, die mehr Möglichkeiten bieten wie z.B. Joysticks oder Controller. Controller haben darüber hinaus oft auch eine taktile Ausgabe, indem sie vibrieren. Es gibt auch Geräte, die eine Eingabe über Drucksensoren ermöglichen (z.B. 3D-Touch Technologie von Apple). Eine experimentelle Technologie, die sehr umfangreiche Ein- und Ausgabe über den Tastsinn ermöglicht, ist z.B. Geomagic Touch.

### **3.6 Geschmacks- und Geruchssinn**

Der Geruchs- und Geschmackssinn wird außerhalb der Forschung in der MCI bisher kaum angesprochen. Der Hauptgrund dafür ist, dass es bisher keinen realistischen Weg gibt, Geruch oder Geschmack zu speichern und im Nachhinein zu reproduzieren. Auch in der Wahrnehmung des Menschen spielen Geruch und Geschmack im Alltag eher eine untergeordnete Rolle.

**Gustatorische Wahrnehmung** Dies ist der Geschmackssinn. Der Geschmackssinn wird stimuliert, indem spezielle Rezeptoren (Geschmacksknospen) auf der Zunge stimuliert werden.

**Olfaktorische Wahrnehmung** Dies ist der Geruchssinn. Hier werden Rezeptoren in der Nase stimuliert. Der Geruchssinn ist für den Menschen ein besonders hilfreicher Sinn, um sich Orte, Situationen, Emotionen oder Identifikationen besser zu merken. Evolutionär dient er außerdem der Gefahrenerkennung.

### **3.7 Psychophysik**

#### **Definition: Psychophysik**

Psychophysik beschäftigt sich mit der Wechselbeziehung zwischen subjektivem Erleben und quantitativ messbaren Reizen.

In der Psychophysik geht es vor allem darum, Formeln und Messwerte für subjektiv wahrgenommene Phänomene zu finden. Ein Beispiel dafür ist das folgende.

### 3.7.1 Weber-Fechner-Gesetz

#### Definition: Weber-Fechner-Gesetz

Das Weber-Fechner-Gesetz beschreibt den Zusammenhang zwischen der Reizstärke  $R$  und der Erlebnisintensität  $E$  durch folgende Formel:

$$E = k \cdot \log R + f$$

Dabei sind  $k$  und  $f$  Konstanten, die bei jedem Menschen unterschiedlich sind.

Die Gültigkeit des Weber-Fechner-Gesetz wurde experimentell mehrfach durch den gerade noch wahrnehmbaren Unterschied (*engl. Just Noticeable Difference*) nachgewiesen. Das Gesetz macht vor allem deutlich, dass die wahrgenommene Intensität logarithmisch von der Reizstärke abhängt.

#### Beispiel: Weber-Fechner-Gesetz

In einem absolut dunklen Raum zündet man eine Kerze an. Der Helligkeitsunterschied ist offensichtlich wahrnehmbar. Zündet man nun eine zweite Kerze an, ist der wahrgenommene Unterschied der Lichtintensität sehr groß (es ist doppelt so hell). Zündet man aber in einem Raum, in dem bereits 1000 Kerzen leuchten, eine weitere an, ist der Helligkeitsunterschied so gering, dass man ihn nicht wahrnimmt. Genau dieses Phänomen beschreibt das Weber-Fechner-Gesetz in allgemeiner Form.

### 3.7.2 Adaption

Die Sensorische Adaption ist ein Effekt, der eintritt, wenn der Mensch über längere Zeit einem *kontinuierlichen* Reiz ausgesetzt ist. Dann passen sich die Rezeptoren entsprechend auf den konstanten Reiz an, sodass er nicht mehr bewusst wahrgenommen wird.

Sensorische Adaption

#### Beispiel: Adaption

Geht man aus dem hellen Sonnenlicht in einen sehr dunklen Keller, kann man zunächst nichts sehen, da die Augen noch an das helle Licht von draußen gewöhnt sind. Nach einiger Zeit gewöhnen sich die Augen aber an die Dunkelheit (Adaption), sodass das Sehen wieder möglich ist.

Sensorische Adaption beschränkt sich nicht nur auf die visuelle Wahrnehmung, sondern findet genauso beim Hören, Riechen, Fühlen und allen anderen Sinnen statt.

### 3.7.3 Habituation

Die Habituation ist ein ähnlicher Effekt wie die Adaption. Er tritt auf, wenn ein Individuum einem Reiz wiederholt ausgesetzt ist, der sich als unbedeutend erwiesen hat. Die

Habituation

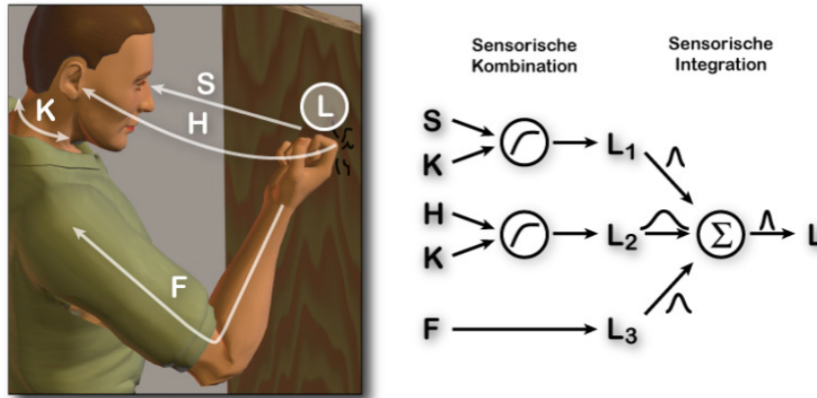


Abbildung 3.13: Sensorische Integration

sensorische Adaption findet, wie der Name andeutet, auf der Ebene der Sensorik statt und kann somit auch nicht unterdrückt werden. Habituation dagegen findet auf einer höheren Ebene statt. Ein Beispiel für Habituation ist das Ticken einer Uhr. Nach einiger Zeit nimmt man die Uhr nicht mehr bewusst wahr. Im Gegensatz zur Adaption kann man sich aber bewusst auf das Ticken konzentrieren und so die Habituation umgehen.

### 3.7.4 Sensorische Integration

Im Kapitel Wahrnehmung haben wir verschiedene Sinne des Menschen kennen gelernt und festgestellt, wie diese mit unserer Wahrnehmung zusammenhängen. Wir haben die Sinne aber bisher hauptsächlich unabhängig voneinander betrachtet. In der Realität nehmen wir kontinuierlich Reize über alle Sinne gleichzeitig auf. Das ist sehr wichtig, da die Informationen eines einzelnen Sinnes oft nicht ausreichen, um ein ausreichendes Bild von der eigenen Umwelt zu erhalten. Die Kombination der einzelnen Sinneseindrücke zu einem Gesamteindruck nennt man Sensorische Kombination bzw. Sensorische Integration.

Sensorische  
Kombination  
Sensorische  
Integration.

**Sensorische Kombination** Im ersten Schritt werden bestimmte Sinne zu einem Zwischeneindruck kombiniert. In Abbildung 3.13 sind das zum einen der visuelle Sinn und die Propriozeption und zum anderen der auditive Sinn und die Propriozeption. An dieser Stelle ist die Propriozeption besonders wichtig, da dadurch aus den visuellen und auditiven Informationen absolute Positiondaten abgeleitet werden können. Dadurch kann aus der relativen Position zum Kopf eine absolute Position zum eigenen Körper abgeleitet werden (die absolute Positionsbestimmung ist natürlich nur so gut wie die Tiefen- bzw. Distanzhinweise es erlauben).

**Sensorische Integration** Im zweiten Schritt werden die Zwischeneindrücke zu einem Gesamteindruck kombiniert. Dabei werden die einzelnen Zwischeneindrücke unterschiedlich gewichtet. Der visuelle Sinn hat hier meistens die höchste Priorität. Danach kommt



normalerweise der auditive Sinn. Darüber hinaus spielt aber auch A-priori Wissen eine wichtige Rolle, durch das beurteilt wird, wie realistisch ein bestimmter Sinneseindruck in der aktuellen Situation ist. Im psychophysischen Sinne kann man dies so ausdrücken: Der Gesamteindruck  $L$  errechnet sich aus den einzelnen Reizen  $L_i$  durch

$$L = \sum_i w_i \cdot L_i \quad \text{mit} \quad \sum_i w_i = 1$$

## 4 Kognition

In diesem Kapitel geht es um das, was umgangssprachlich als Denken bezeichnet wird. Im Model Human Prozessor befinden wir uns hier um kognitiven Prozessor. Im Gehirn befindet sich der Teil der Kognition, mit dem wir uns beschäftigen im vorderen Teil. In diesem Kapitel geht unter Anderem es darum, wie das Gedächtnis funktioniert, wie wir lernen, logisch denken und wie Entscheidungsprozesse ablaufen können.

Für die Mensch-Computer-Interaction ist es wichtig, die Kognition zu verstehen, da so Benutzerschnittstellen so optimiert werden können, dass man sie sofort versteht.

### 4.1 Gedächtnis

Schon im Bereich der Wahrnehmung haben wir festgestellt, dass uns ein ständiger Strom an Daten erreicht. Das Gehirn kann diese Datenmengen weder verarbeiten noch speichern. Daher findet Filterung und Selektion statt (s. Abschnitt 2.3.1). Bevor die Informationen dann im Gedächtnis gespeichert werden, werden sie noch im Gehirn verarbeitet und dabei gegebenenfalls transformiert. Die Verarbeitung von Informationen betrachten wir detailliert im Abschnitt ?? Die Datenmengen, die tatsächlich im Gedächtnis gespeichert werden können, sind in Tabelle 4.1 dargestellt.

	pro Stunde	pro Jahr	im Leben
Lesen	50 kB	427 MB	33 GB
Hören	50 MB	427 GB	33 TB
Sehen	1 GB	9 TB	670 TB

Tabelle 4.1: Speicherbedarf für die Informationsverarbeitung

#### Definition: Multi-Speicher-Modell

Das Multi-Speicher-Modell ist ein Gedächtnismodell, das drei miteinander verbundene Speicher unterscheidet:

1. Sensorische Register *Ultrakurzzeitgedächtnis*
2. Arbeits-/Kurzzeitgedächtnis *engl. Short-Term Memory, STM*
3. Langzeitgedächtnis *engl. Long-Term Memory, LTM*

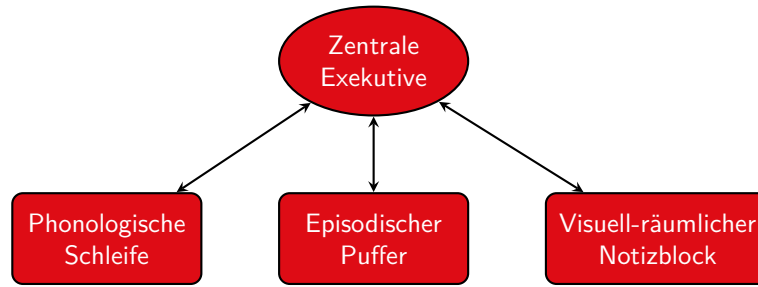


Abbildung 4.1: Arbeitsgedächtnismodell

### 4.1.1 Sensorische Register

Die Sensorischen Register bezeichnet man umgangssprachlich auch als Ultrakurzzeitgedächtnis. Diese Register dienen als sehr flüchtiger Zwischenspeicher für Stimuli. Man unterscheidet die Register nach den einzelnen Sinnen:

Sensorische Register

1. Ikonische Register
2. Echonische Register
3. Haptische Register

Die Register haben nur eine sehr kurze Speicherdauer (ca. 0,5 s-1 s für visuelle Informationen, bzw. 4 s-5 s für auditive Informationen).

Aus den sensorischen Registern gelangen die Informationen ins Arbeitsgedächtnis, indem man Aufmerksamkeit auf einen Stimulus lenkt. Dadurch werden dann die Elemente aus den Registern extrahiert. Nachdem 4-10 Elemente extrahiert wurden, werden die Informationen aus den Registern nach dem FIFO-Prinzip (first-in-first-out) gelöscht.

### 4.1.2 Arbeitsgedächtnis

Das Arbeitsgedächtnis ist der im Alltag am stärksten beanspruchte Speicher im Gehirn. Es ist vergleichbar mit dem Arbeitsspeicher bei einem Computer. Hier werden Zwischenergebnisse abgelegt, die in kürze weiterverwendet werden. Das Arbeitsgedächtnis braucht man für die meisten alltäglichen Aufgaben wie Lesen, Rechnen, Schreiben, Sprechen, Zuhören usw. Entsprechend ist die Zugriffszeit auf dieses Gedächtnis mit weniger als 70 ms auch sehr schnell. Im Arbeitsgedächtnis bleiben Informationen für etwa 15 s-30 s erhalten.

Experimente haben gezeigt, dass das Arbeitsgedächtnis aus mehreren Teilen besteht. Versucht man parallel zwei Rechenaufgaben parallel zu lösen, fällt dies den meisten Menschen relativ schwer. Wird man aber während des Rechnens von einer anderen Person angesprochen, schaffen es die meisten Menschen sehr gut, sowohl die Rechenaufgabe zu bearbeiten als auch dem Gesprächspartner zuzuhören. Draus ergibt sich für das Arbeitsgedächtnis eine Aufteilung in mehrere Subspeicher (Abbildung 4.1).

Die einzelnen Gedächtnisteile haben je eigenene Zuständigkeiten:

**Phonologische Schleife** ist dafür Zuständig Sprache zu verarbeiten.

**Episodischer Puffer** Speicher Erlebnisse und Erfahrungen.

**Visuell-räumlicher Notizblock** Interpretiert visuelle Informationen.

Bis zu einem gewissen Grad können die drei Gedächtnisteile auch miteinander interagieren, allerdings werden solche Aufgaben für den Menschen schnell sehr schwierig.

**Chunking** Das Arbeitsgedächtnis hat nur eine sehr begrenzte Kapazität. Insgesamt kann es weniger als 10 Informationen gleichzeitig beinhalten. Das reicht im Alltag normalerweise nicht aus. Daher werden mehrere Elemente im Arbeitsgedächtnis zu einer Sinneseinheit von Informationen (*engl. Chunk*) gruppiert, um die Kapazität künstlich zu vergrößern. Je nach Quelle können zwischen 4 und 9 solcher Chunks im Arbeitsgedächtnis gespeichert werden.

#### Beispiel: Chunking

Chunking kann man sehr gut beobachten, wenn man versucht, sich Telefonnummern zu merken. Viele Menschen merken sich nicht jede einzelne Ziffer, sondern 2-3 stellige Zahlen. Hier werden also auch einzelne Ziffern zu einer Einheit gruppiert, um den benötigten Speicher zu minimieren.

**Primär- und Rezenzeffekt** Diese Effekte bezeichnen das Phänomen, dass man sich an früher, bzw. später eingehende Informationen besser erinnern kann. Wenn man sich beispielsweise eine Ziffernfolge merken soll, fällt es meistens deutlich leichter, sich an die erste (Primäreffekt) bzw. letzte Ziffer (Rezenzeffekt) zu erinnern als an die in der Mitte.

Aus dem Arbeitsgedächtnis gelangen einige Informationen über den Prozess des Memorierens ins Langzeitgedächtnis. Memorieren besteht im wesentlichen darin, eine Information im Arbeitsgedächtnis oft zu transformieren und zu wiederholen. Informationen, die man sich länger als 30s merkt, befinden sich im Langzeitgedächtnis. Aus dem Kurzzeitgedächtnis werden Informationen (Chunks) gelöscht, wenn die damit assoziierte Aufgabe (z.B. eine Rechenaufgabe) abgeschlossen wurde, spätestens aber nach 30s.

Memorieren

### 4.1.3 Langzeitgedächtnis

Das Langzeitgedächtnis ist der Hauptspeicher des menschlichen Gehirns. Informationen, die darin abgelegt werden, können zwischen einigen Minuten bis zu vielen Jahren wieder abgerufen werden. Der Nachteil an der hohen Kapazität ist allerdings, dass die Zugriffszeit mit etwa 0,1s relativ hoch ist. Im Langzeitgedächtnis können alle möglichen Arten von Information gespeichert werden (Fakten, Daten, Töne, Gerüche, Bilder, ...).

Die verschiedenen Arten von Informationen sind im Langzeitgedächtnis in unterschiedlichen Gedächtnisteilen gespeichert. Man unterscheidet dabei das deklarative Gedächtnis und das nicht-deklarative Gedächtnis. Das deklarative Gedächtnis ist besonders wichtig für das bewusste Anwenden von Regeln und für Schlussfolgerungen, während das

Deklaratives  
Gedächtnis

nicht-deklarative Gedächtnis vor allem bei unbewussten Prozessen zum Einsatz kommt. Daher bezeichnet man das deklarative Gedächtnis auch oft als Explizites Gedächtnis und das nicht-deklarative als Implizites Gedächtnis. Insgesamt unterscheiden wir vier verschiedene Teile des Langzeitgedächtnisses:

Explizites Gedächtnis

#### **Episodisches Gedächtnis deklarativ**

Hier werden Erlebnisse und Erfahrungen gespeichert.

#### **Semantisches Gedächtnis deklarativ**

Dieser Gedächtnisteil enthält Faktenwissen.

#### **Prozedurales Gedächtnis nicht-deklarativ**

Hier werden kognitive oder motorische Fertigkeiten gespeichert.

#### **Priming und Konditionierung nicht-deklarativ**

Unbewusste Voreingenommenheit durch vorangegangene Reize.

**Semantisches Netz** Das semantische Netz ist ein formales Modell von Begriffen und ihren Beziehungen zur Wissensrepräsentation im semantischen Gedächtnis. Diese Beziehungen (oder Assoziationen) erlauben es, Begriffe auf semantischer Ebene miteinander zu verbinden. Dadurch entsteht ein Netz von Begriffen, das als semantisches Netz oder auch Assoziativer Speicher bezeichnet wird.

Semantisches Netz

Assoziativer Speicher

**Prozedurales Gedächtnis** In diesem Gedächtnisteil werden Bewegungs- und Denkabläufe gespeichert. Solche kognitive oder motorischen Fertigkeiten lassen sich durch Übung verbessern und beschleunigen. Das umfasst sowohl sportliche (motorische) Bewegungsabläufe, als auch kognitive (z.B. Rechnen).

**Priming** Priming ist ein Effekt, bei dem ein vorangegangener Reiz mein aktuelles Handeln beeinflusst. Geht es zum Beispiel in einem Gespräch um Geld, sogt Priming dafür, dass man bei dem Begriff „Bank“ an ein Geldinstitut denkt. Ginge es in dem Gespräch um Möbel, wäre die erste Assiziation bei „Bank“ vermutlich eher ein Sitzmöbel.

#### **Erinnern vs. Erkennen**

Das Langzeitgedächtnis ist ein Speicher, auf dem nicht gut gearbeitet werden kann. Um die abgespeicherten Informationen wieder abzurufen und anzuwenden, müssen die Daten zuerst aus dem Langzeitgedächtnis zurück in das Arbeitsgedächtnis gebracht werden. Dazu gibt es zwei Möglichkeiten: Erinnern (*engl. Recall*) und Wiedererkennen (*engl. Recognition*). Beim Erinnern kann man noch weiter zwischen Freiem und seriellem Erinnern (ohne bzw. mit Reihenfolge) unterscheiden.

Im Alltag fällt einem auf, dass der Prozess des Erinnerns offenbar deutlich schwieriger ist, als Dinge wiederzuerkennen. Aus diesem Grund haben sich in der MCI auch grafische Benutzerschnittstellen gegenüber Kommandozeilen-Interfaces durchgesetzt (s. Kpitel 7).

Auf modernen Betriebssystemen können Benutzer so ein Programm an seinem Icon wiedererkennen und müssen sich nicht an den Namen und die Funktion des Programmes erinnern.

## Recognize-Act-Zyklus

### Definition: Recognize-Act-Zyklus

Der Recognize-Act-Zyklus ist der elementare kognitive Prozess, bei dem Informationen aus dem Langzeitgedächtnis abgerufen werden (*engl. recognize*), um neue Einheiten im Arbeitsgedächtnis verfügbar zu machen oder zu verändern (*engl. (act)*).

## Vergessen

Der Prozess des Vergessens ist einer der ungeklärten Bereiche der Hirnforschung. Es gibt zwei sehr populäre Theorien, von sich bisher keine über die andere durchsetzen konnte:

**Spurenverfallstheorie** Diese Theorie geht davon aus, dass die Fähigkeit der Erinnerung und des Wiedererkennens mit der Zeit verblasst. Dabei ist nicht alleine die Zeit, sondern besonders die Häufigkeit der Aktivierung relevant. Je häufiger auf ein bestimmtes Element im Langzeitgedächtnis zugegriffen wird, desto später verfällt es und desto länger bleibt es verfügbar. Aus dieser Theorie geht unter Anderem hervor, dass man sich mehr Informationen merken kann, wenn die Lerneinheiten über die Zeit verteilt sind.

**Interferenztheorie** Die Interferenztheorie geht davon aus, dass man Wissen grundsätzlich nicht vergisst, sondern dass Elemente im Gedächtnis überschrieben werden. Man unterscheidet Proaktive und Retroaktive Interferenz/Hemmung. Bei Proaktiver Interferenz beeinträchtigt existierendes Wissen später hinzugefügtes Wissen. Bei Retroaktiver Interferenz wird altes Wissen durch neues Wissen unwiederrufbar überschrieben.

### Beispiel

Die Umstellung von DM zum Euro kann als Beispiel für die proaktive Interferenz angeführt werden. Hier beeinflusst das Wissen um die Währung DM die Denkweise im Euro-System. Ein Beispiel für retroaktive Interferenz kann man Passwörter anführen, die man vergisst, wenn man ein neues Passwort wählt.

## 4.2 Lernen

### Definition: Lernen

Lernen beschreibt den absichtlichen und beiläufigen Erwerb von kognitiven und motorischen Fähigkeiten und Fertigkeiten. Lernen ist außerdem der Prozess der Veränderung des Verhaltens, Denkens oder Fühlens aufgrund von Erfahrung oder neu gewonnenen Einsichten.

### 4.2.1 Lernmethoden

Das Ziel einer Lernmethode ist es, dem Menschen das Lernen zu erleichtern. Aus dem Alltag wissen die meisten Menschen, dass das Lernen von Fakten (also das Lernen von semantischem Wissen) besonders schwierig ist. Daher haben die meisten Lernmethoden das Ziel, dass man Faktenwissen besser lernen kann. Das funktioniert normalerweise so, dass durch die Lernmethode das Faktenwissen nicht allein im semantischen Gedächtnis abgespeichert wird, sondern auch im prozeduralen oder episodischen Gedächtnis. Es gibt viele verschiedene Lernmethoden. Wir betrachten hier nur einige wenige:

**Loci-Methode** Die Loci-Methode zählt zu der Kategorie der Gedächtnisstützen. Dabei werden die Lerninhalte in eine fiktive Struktur eingegliedert. Möchte man sich zum Beispiel eine lange Zahlenkette merken, reicht es nicht die Eingabe nur zu Chunks (vgl. Abschnitt ??). Nach der Loci-Methode versucht man nun die einzelnen Chunks (zum Beispiel 2-3 stellige Zahlen) ins episodische Gedächtnis zu übertragen, indem man mit jeder Zahl ein Wort assoziiert und so eine Geschichte erhält. Die Voraussetzung hierfür ist allerdings, dass man bereits vorher das Mapping zwischen Chunks und Wörtern auswendig gelernt hat.

**Gedächtnispalast** Bei dieser Gedächtnisstütze versucht man Fakten in eine räumliche Struktur einzubetten. Anstatt sich also rein semantische Fakten zu merken, assoziiert man diese mit einem Ort in seinem Gedächtnispalast. Bei Abrufen der Daten geht man dann an den entsprechenden Ort im Palast und kehrt so die Assoziation um, um wieder die Fakten zu erhalten.

**Prozedurale Methoden** Die obigen beiden Lernmethoden arbeiten vorwiegend mit dem episodischen Gedächtnis. Gerade im Bereich MCI ist aber auch das prozedurale Gedächtnis eine gute Gedächtnisstütze. Zum Beispiel fällt es vielen Menschen schwer, sich ein Passwort zu merken. Allerdings prägen sich die Fingerbewegungen, die zum Eintippen des Passwortes nötig sind, sehr schnell ein. So wird das Passwort als rein semantisches Wissen im prozeduralen Gedächtnis gespeichert.

### 4.2.2 Fertigkeitserwerb

Der Fertigkeitserwerb findet hauptsächlich im prozeduralen Gedächtnis statt. Der Prozess des Lernen unterscheidet sich dabei nicht wesentlich bei motorischen und kognitiven Fähigkeiten. Er lässt sich in drei Phasen einteilen:

**1. Kognitive Phase** In der ersten Phase ist es noch sehr schwer die Fähigkeit anzuwenden. Man muss viel Denkleistung aufwenden, um eine entsprechende Aufgabe zu erfüllen.

Lernt man zum Beispiel eine bestimmte Schlagtechnik im Tennis, muss man bei jedem Schlag nachdenken, wie der entsprechende Bewegungsablauf funktioniert.

**2. Assoziative Phase** In der zweiten Phase, funktioniert die Fähigkeit schon zum Teil unterbewusst. Hier kommt es nun darauf an, viel zu trainieren, um die Schnelligkeit und Genauigkeit der Fähigkeit zu verbessern.

**3. Autonome Phase** Nach viel Training beginnt die Autonome Phase. In dieser Phase läuft die Fähigkeit quasi völlig automatisiert ab. Es erfordert kaum noch aktive Denkleistung und geht sehr schnell. Die Fertigkeiten Lesen oder Sprechen fallen in diese Kategorie.

### 4.2.3 Lernform

#### Definition: Lernform

Eine Lernform ist ein methodisch-didaktischer Ansatz zum Wissens- und Fertigkeitserwerb und dessen Umsetzung in unterschiedlichen Kontexten.

Wir werden hier nur ein Beispiel für Lernformen betrachten: die Dekomposition. Dekomposition ist die Zerlegung einer größeren Aufgabe in kleinere Teilaufgaben. Das ist für die MCI besonders interessant, da viele Benutzerinteraktionen durch Dekomposition vereinfacht werden können. Möchte man zum Beispiel eine Website öffnen, kann man dies zerlegen in „Webbrowser öffnen“, „URL eintippen“, „Website öffnen“. Jeder dieser Schritte kann auch noch weiter zerlegt werden (siehe dazu Abschnitt ??).

Dekomposition

### 4.2.4 Lernbewertung

In der MCI gibt es sehr häufig die Situation, dass der Benutzer eine neue Vorgehensweise lernen muss. Entwickelt man eine neue Technologie, muss der Benutzer lernen, wie er damit umgeht. Nun gibt es viele verschiedene Möglichkeiten, dem Benutzer zu erklären, wie die Technologie funktioniert (z.B. Tutorials, Videos, Vorführungen). Daher ist es wichtig, ein Verfahren zu haben, mit dem man beurteilen kann, welche der möglichen Lernformen am besten funktioniert. Dazu geht man folgendermaßen vor:

1. Man wählt eine Transfergruppe, die mit dem neuen Verfahren lernt.
2. Man wählt eine Kontrollgruppe, die ohne das neue Verfahren lernt.
3. Man wählt eine Metrik, nach der die Güte des Verfahrens beurteilt werden soll (z.B. die benötigte Zeit oder die Genauigkeit)
4. Man lässt die beiden Gruppen unabhängig voneinander die eine Aufgabe bearbeiten.

Transfergruppe

Kontrollgruppe



In der Auswertung gibt es nun zwei Werte, die interessant sind:

$$\text{Leistung} = \frac{\text{Zeit}_{\text{Kontrollgruppe}} - \text{Zeit}_{\text{Transfergruppe}}}{\text{Zeit}_{\text{Kontrollgruppe}}} \cdot 100$$

Bei der Leistung fragt man also „Wie viel % der Anwendungszeit wird eingespart?“

$$\text{Effektivität} = \frac{\text{Zeit}_{\text{Kontrollgruppe}} - \text{Zeit}_{\text{Transfergruppe}}}{\text{Zeit}_{\text{neue Lerntechnik}}}$$

Hier fragt man sich: „Wie ist das Verhältnis vom Zeitgewinn zum Lernaufwand?“

#### Aufgabe 12: Lernbewertung

Angenommen es gibt zwei verschiedene Methoden das 10-Finger-System zu erlernen, um einen Text abzutippen.

**Methode 1** Die Transfergruppe lernt 5 min und braucht 10 min zum Abtippen des Textes.

**Methode 1** Die Transfergruppe lernt 10 min und braucht 5 min zum Abtippen des Textes.

**Methode 1** Die Kontrollgruppe braucht 15 min zum Abtippen des Textes.

Welche Methode führt zur höchsten Leistung?

### 4.2.5 Potenzgesetz der Übung

Das Potenzgesetz der Übung beschreibt, wie sich die Ausführungszeit einer *motorischen* Aufgabe durch Übung reduziert. In der allgemeinen Form ist dies

$$T_n = T_1 \cdot n^{-a}$$

wobei  $T_1$  die Aufgabenschwierigkeit,  $a$  der Lernparameter und  $T_n$  die benötigte Zeit für die Aufgabe nach  $n$  Wiederholungen ist.

Was aus dem Potenzgesetz hervorgeht, ist vor allem, dass man häufig gerade am Anfang besonders gut lernt. Je öfter man aber eine Aufgabe wiederholt hat, desto geringer ist der Ertrag der einzelnen Wiederholung. Damit lässt sich auch die 80-20 Regel des Lernens gut begründen (Man lernt 20 % der Zeit für 80 % des Ergebnisses, und die restlichen 80 % der Zeit für die fehlenden 20 %). Diese Regel bezieht sich genau darauf, dass man gerade am Anfang sehr schnell lernt, später aber sehr viel lernen muss, um einen merklichen Ertrag zu erzielen.

### 4.3 Logisches Denken

Im Gebiet der Logik untersucht man die Struktur von Argumenten in Hinblick auf ihre Gültigkeit, unabhängig vom Inhalt der Aussagen. Die Logik hat ihren Ursprung in der

Philosophie und hat an sich nicht viel mit der Funktionsweise von logischem Denken zu tun. Allerdings sind die grundlegenden Mechanismen der Logik auch im Logischen Denken des Menschen wiederzufinden. Damit können wir folgende Definition aufstellen:

**Definition: Logisches Denken**

Das logische Denken beinhaltet alle Vorgänge, die aus der inneren Beschäftigung mit Vorstellungen, Erinnerungen und Begriffen versuchen, Erkenntnis zu formen. Das logische Denken ist die kognitive Fähigkeit des Menschen.

Das logische Denken des Menschen wird unterteilt in drei Arten:

**Deduktion** geht vom Allgemeinen zum Besonderen.

**Induktion** geht vom Einzelnen zum Allgemeinen.

**Abduktion** schließt vom Resultat und der Regel auf den Fall.

Die drei Arten wurden von Charles Sanders Peirce folgendermaßen charakterisiert:

*„Deduction proves that something must be; Induction shows that something actually is operative; Abduction merely suggests that something may be.“*

— Charles Sanders Peirce

### 4.3.1 Deduktion

Das deduktive Schließen verwendet logische Regeln, um Folgerungen zu schließen. Eine gültige Deduktion zeichnet sich dadurch aus, dass sich aus der Regel die entsprechende logische Konsequenz ableiten lässt. Gegeben ist zum Beispiel die **logische Regel** „Alle Vielfachen von 15 sind durch 3 teilbar“. Dann besteht eine gültige Deduktion aus zwei Schritten:

**Bedingung** Die Zahl 45 ist ein Vielfaches von 15.

**Logische Konsequenz** Die Zahl 45 ist durch 3 teilbar.

Hier wurde also die Regel angewendet, wodurch eine Konsequenz entstanden ist.

#### Aufgabe 13

Sind die folgenden Deduktionen korrekt?

1. Am Sonntag scheint die Sonne.
  - a) Es ist Sonntag. Also scheint die Sonne.
  - b) Die Sonne scheint. Also ist Sonntag.
2. An manchen Montagen scheint die Sonne.

- a) Es ist Montag. Also scheint die Sonne.
  - b) Es scheint die Sonne. Also ist Montag.
3. Alle geraden Zahlen sind durch 2 teilbar.
- a) Die Zahl 10 ist gerade. Also ist 10 durch 2 teilbar.
  - b) Die Zahl 5 ist nicht gerade. Also ist 5 auch nicht durch 2 teilbar.
  - c) Die Zahl 4 ist durch 2 teilbar. Also ist 4 gerade.

### 4.3.2 Induktion

Bei der Induktion ist das Vorgehen genau anders herum als bei der Deduktion. Hier wird von einzelnen beobachteten Fällen auf allgemeine Regeln generalisiert.

#### Beispiel: Induktion

Jeder Elefant, den ich bisher gesehen habe, hat einen Rüssel. Also hat jeder Elefant einen Rüssel.

Induktion ist mathematisch-formal nicht korrekt. Bei der Induktion ist immer eine gewisse Unsicherheit vorprogrammiert, da es immer sein kann, dass die Beobachtungen, die man gemacht hat, zufällig gerade Spezialfälle waren, sodass die induzierte Regel falsch ist. Das ist aber nicht sonderlich schlimm, da es sehr einfach ist, falsche Induktionen zu falsifizieren. Im obigen Beispiel würde die Beobachtung eines einzigen Elefanten ohne Rüssel ausreichen, um die Induktion zu falsifizieren.

### 4.3.3 Abduktion

Die Abduktion ist ein logischer Vorgang, in dem zu einem Phänomen eine erklärende Hypothese gebildet wird. Hier beobachtet man also ein Resultat und kennt bereits eine Regel, nach der eben dieses Resultat folgen würde. Jetzt wendet man diese Regel rückwärts an, und erhält somit eine Hypothese, wie das Resultat entstanden sein könnte.

#### Beispiel

- Logisch Regel: Falls A wahr wäre, würde C folgen.
- Resultat: Tatsache C.
- Abduktion: A ist wahr.

Offensichtlich ist auch die Abduktion nicht formal korrekt, da außer der Ursache A noch viele weitere Ursachen möglich wären, hier aber außer Acht gelassen werden.

Damit Abduktion sinnvoll angewendet werden kann, ist vor allem viel Lebenserfahrung notwendig. Durch die Lebenserfahrung lernt man viele mögliche Regeln kennen und kann dadurch deutlich realistischere Hypothesen bilden.

## 4.4 Entscheidungen treffen (Gesetz von Hick)

Entscheidungsprozesse verwerten Informationen aus dem Langzeitgedächtnis und logisches Denken, um zwischen mehreren verfügbaren Optionen auszuwählen. Entscheidungsprozesse finden jeden Tag bewusst und unbewusst statt. Auch bei der Interaktion des Menschen mit dem Computer, müssen ständig Entscheidungen getroffen werden. Diese umfassen zum Beispiel die Wahl des Eingabegerätes (Tastatur oder Maus) oder die Entscheidung zwischen Optionen auf dem Bildschirm. Uns interessiert nun vor allem, wie lange es dauert, sich zwischen einigen Möglichkeiten zu entscheiden und wie wir durch intelligentes Schnittstellendesign diese Zeiten möglichst verkürzen können.

Genau mit dieser Frage hat sich auch W.E. Hick beschäftigt und hat das Gesetz von Hick aufgestellt:

### Definition: Gesetz von Hick

Die Zeit  $T$  (in Sekunden), die eine Person benötigt, um eine Entscheidung aus einer Menge mit  $n$  **bekanntem, gleich wahrscheinlichen** Elementen zu treffen, errechnet sich durch

$$T = b \cdot \log_2(n + 1)$$

Dabei ist  $b$  eine empirisch bestimmte Konstante, die sich je nach Kontext und von Person zu Person unterscheidet.

Häufig wird anstelle von  $b$  auch der *Processing Speed* verwendet. Hier gilt dann,  $\text{Processing Speed} = \frac{1}{b}$  und damit:

$$T = \frac{\log_2(n + 1)}{\text{Processing Speed}}$$

Untersuchungen haben einerseits ergeben, dass es eine direkte Korrelation zwischen IQ und *Processing Speed* gibt und dass es andererseits möglich ist, den *Processing Speed* durch viel Übung zu steigern.

Wie schon bei der perzeptuellen Ebene beim Weber-Fechner-Gesetz (Abschnitt 3.7.1) ist auch hier ein logarithmischer Zusammenhang zu finden.

Damit das Gesetz von Hick gilt, ist wichtig, dass alle Elemente bereits im Voraus bekannt sind. Ist das nicht der Fall, braucht man mindestens zusätzlich die Zeit, um alle verfügbaren Elemente *kennen zu lernen*. Außerdem müssen die verfügbaren Elemente vom Menschen zumindest als gleich wahrscheinlich wahrgenommen werden. Ist eine Option völlig abwegig oder fällt dem Menschen eine Möglichkeit nicht ein, verringert sich die Anzahl der Elemente aus denen entschieden werden muss, entsprechend.

### Beispiel

Am Computer möchte man den aktuell markierten Text kopieren. Hier hat man nun die Wahl zwischen drei Optionen:

**Kontextmenü** Über einen Rechtsklick die Option „Kopieren“ auswählen.

**Tastenkürzel** Über ein Tastenkommando kopieren.

**Programmmenü** Im Menü „Bearbeiten“ die Option „Kopieren“ auswählen.

Die Entscheidung zwischen diesen drei Optionen braucht dann nach dem Gesetz von Hick

$$T = b \cdot \log_2(3 + 1) = 2b$$

Wichtig ist hier, dass die Zeit, die das Ausführen jeder Aktion braucht, noch nicht mit berücksichtigt wird. Die Zeit  $2b$  braucht alleine der Entscheidungsprozess. Es gilt dann also

$$\text{Reaktionszeit} = \text{Bewegungszeit} + \frac{\log_2(n + 1)}{\text{Processing Speed}}$$

#### Aufgabe 14: Gesetz von Hick

Aufgabe zur Menu-Auswahl

Das Ergebnis aus der obigen Aufgabe lässt sich verallgemeinern: Im Allgemeinen kostet eine Auswahl aus einer komplexen Struktur von Optionen mehr Zeit, als die Auswahl bei einfachen Alternativen. Auch bei einer großen Menge von Optionen ist der Entscheidungsprozess nach dem Gesetz von Hick bei einer sehr großen Liste schneller als bei verschachtelten Optionen. Man muss ab einer gewissen Menge Optionen allerdings berücksichtigen, dass für einen Entscheidungsprozess nach Hick die verfügbaren Optionen im Kurzzeitgedächtnis gespeichert werden müssen. Da dessen Kapazität begrenzt ist und häufig auch der Platz auf dem Bildschirm nicht unendlich groß ist, kann eine verschachtelte Auswahl in manchen Situationen doch angebracht sein.

## 4.5 Aufmerksamkeit und Belastung

Jeder hat schon einmal erlebt, dass einem ein Detail entgangen ist, auf das man nicht geachtet hat. Es ist genauso bekannt, dass man, wenn man müde ist, deutlich mehr Flüchtigkeitsfehler macht. Dies motiviert die Definition von Aufmerksamkeit und Belastung.

### 4.5.1 Aufmerksamkeit

#### **Definition: Aufmerksamkeit**

Aufmerksamkeit ist die Zuweisung von Bewusstseinsressourcen auf Inhalte bzw. Aufgaben. **Konzentration** ist ein Maß für die Intensität und Dauer der Aufmerksamkeit.

Wie wichtig Aufmerksamkeit ist, wenn es darum geht, Details wahrzunehmen, zeigt der Cocktail-Party-Effekt. Dieser Effekt beschreibt die Fähigkeit des Menschen, auch

Cocktail-Party-Effekt

in dem Stimmengewirr einer Cocktail-Party eine geregelte Unterhaltung zu führen und seine Gesprächspartner dabei auch zu verstehen. Untersuchungen hierzu haben ergeben, dass in so einer Situation das Stimmengewirr um einen herum um bis zu 15 dB gedämpft wird.

Der Cocktail-Party-Effekt tritt allerdings nur ein, wenn die Aufmerksamkeit bereits auf einem Gespräch ist. Um die Aufmerksamkeit auf ein Gespräch oder eine andere Sache zu lenken, gibt es einige Fokussierungsereignisse:

- Besonders starke Reize (z.B. ein Lichtblitz oder ein Knall)
- Bewegung
- Farbe und Kontrast
- ...

#### 4.5.2 Belastung

Die Menge der Bewusstseinsressourcen, die durch Aufmerksamkeit beansprucht werden können, ist begrenzt. Das heißt, dass durch das Bearbeiten einer Aufgabe (wie zum Beispiel dem Zuhören bei einer Cocktail-Party) weniger kognitive Ressourcen für andere Aufgaben übrig bleiben. Das wird vor Allem bei Aufgaben spürbar, die unter Zeitdruck gelöst werden müssen oder bei denen eine starke emotionale Belastung eintritt.

Bei jeder Aufgabe, die man bearbeitet, finden immer drei unterschiedliche Arten von kognitiver Belastung statt. Die Anteile der drei Teile an der Gesamtbelastung unterscheiden sich je nach Aufgabe:

**Lernbezogene Belastung** Diese Art von Belastung tritt zum Beispiel auf, wenn man lernt, eine Software zu bedienen. Dies ist die Belastung, die tatsächlich notwendig ist, um die gestellte Aufgabe zu erfüllen.

**Intrinsische Belastung** Diese Belastung ist mit der Aufgabe selbst verbunden. Hierzu zählt die Überlegung, welche Aktionen als nächstes ausgeführt werden müssen oder welche Wörter man zum Schreiben eines Textes wählt.

**Extrinsische Belastung** Diese Belastung entsteht durch das verwendete Medium. Hierzu zählen z.B. Schwierigkeiten bei der Bedienung der Benutzerschnittstelle.

Wo die kognitive Belastungsgrenze ist, unterscheidet sich von Mensch zu Mensch, hängt aber auch von der Tagesform ab. Ist man zum Beispiel müde, ist die Belastungsgrenze auch deutlich niedriger. Wenn die Belastung zu hoch wird, zeigt sich dies darin, dass man beim Bearbeiten einer Aufgabe extrem viele Fehler machen oder es nicht schaffen mehrere Aufgaben parallel zu bearbeiten, wodurch eine Aufgabe abgebrochen werden muss.

**Messen kognitiver Belastung** Kognitive Belastung wird gemessen, indem Probanden eine Hauptaufgabe und mehrere Nebenaufgaben gestellt werden. Die Hauptaufgabe ist dabei normalerweise deutlich anspruchsvoller als die Nebenaufgaben. Die Messung erfolgt dann anhand der Nebenaufgaben: Je weniger davon bearbeitet werden können, bzw. je mehr Fehler gemacht werden, desto schwieriger war offenbar die Hauptaufgabe (da diese mehr kognitive Ressourcen beansprucht hat).

# 5 Kommunikation und Handlung

In diesem Kapitel geht es um den motorischen Prozessor im Human-Model-Processor. Nachdem im Kognitiven Prozessor die Eingabe des perzeptuellen Prozessors verarbeitet wurde, wird jetzt die Ausgabe generiert. Die Ausgabe kann viele unterschiedliche Formen annehmen (z.B. Sprache oder Bewegung). Im Gehirn finden diese Prozesse entweder oben oder an der Seite statt. In diesem Kapitel werden wir einige Arten der Motorik kennen lernen und sehen, wie sich Denkfehler im motorischen Prozessor manifestieren. Zuletzt lernen wir einige Modelle aus der Mensch-Computer-Interaktion kennen, mit denen die Motorik des Menschen als Eingabe für einen Computer modelliert wird.

## 5.1 Kommunikation

Kommunikation erfolgt im klassischen Sinne zwischen zwei oder mehr Menschen. Ein Großteil davon findet über natürliche Sprache statt, aber auch Mimik, Stimmlage und Gestik spielen eine wichtige Rolle. In der Mensch-Computer-Interaktion interessiert uns vor allem die Kommunikation zwischen Mensch und Computer. Hier stellen wir zuerst fest, dass der Computer im Gegensatz zum Menschen deutlich stärker eingeschränkt ist. Der Computer muss außerdem alle Daten irgendwie Codieren, um damit umgehen zu können (als Text, als Grafik, ...). Hier spricht man von Multicodalität.

Multicodalität

Analog zu den verschiedenen Wegen, Informationen zu codieren, gibt es auch verschiedene Wege der Informationsvermittlung über die Mensch-Computer-Schnittstelle. Wie im Abschnitt „Wahrnehmung“ bereits angedeutet wurde, erfolgt der größte Teil der Kommunikation über den visuellen bzw. auditiven Kanal. Der Computer stellt also Informationen auf einem grafischen Display oder über die Kopfhörer dar, sodass diese vom Benutzer wahrgenommen werden können.

Zuletzt muss der Mensch die vom Computer gesendeten Informationen aufnehmen. Dies erfolgt über verschiedene Modalitäten (Auge, Ohr, ...).

Modalität

Zur modellierung der Kommunikation zwischen Mensch und Computer über eine Schnittstelle werden wir die beiden folgenden Paradigmen betrachten:

**Mensch-Computer-Kommunikation** Der Mensch sieht den Computer als Kommunikationspartner.

**Mensch-Computer-Interaktion** Der Mensch sieht den Computer als Handlungsraum, mit dem er interagieren kann.

Es gibt neben diesen beiden Paradigmen auch noch weitere Sichten auf die Kommunikation zwischen Mensch und Computer, mit denen wir uns aber nicht weiter beschäftigen werden:



**Computer als Arbeitsmittel oder Werkzeug** Der Mensch setzt den Computer ein oder bedient ihn.

**Computer als Medium** Der Computer ist ein Mittler zur Anwendungswelt. Der Mensch interagiert mit dieser Welt über den Computer als Schnittstelle.

**Computer als künstliche Realität** Der Mensch bewegt sich in einer vom Computer geschaffenen, künstlichen Welt.

### 5.1.1 Kommunikationssicht

Forschungen haben gezeigt, dass jedes Objekt, das dem Menschen sprachlich gegenübertritt, grundsätzlich zunächst als Kommunikationspartner wahrgenommen wird. Ein relativ frühes, bekannte Beispiel aus der MCI hierfür ist das „Eliza“-Programm oder „Cleverbot“. Inzwischen gibt es auch Software, die natürliche Sprache erkennt und eine beidseitig verbale Kommunikation erlaubt (z.B. Siri).

Aus dem Fokus auf der Sprachlichen Interaktion geht hervor, dass Modelle auf Grundlage der Kommunikationssicht sehr deskriptiv sind (also viel mit Text bzw. Sprache arbeiten). Das heißt auch, dass hier deutlich mehr Erinnern im Gegensatz zu Wiedererkennen (vgl. Abschnitt 4.1.3) nötig ist.

Der Mensch nimmt damit den Computer als natürlichen Kommunikationspartner wahr, mit dem man sich unterhalten kann. Dasselbe Konzept kann jedoch auch auf andere Modalitäten wie zum Beispiel die visuelle Ebene übertragen werden. Insbesondere durch die sich immer weiter verbreitenden VR-Technologien wird die Idee der Kommunikation mit dem Computer in einer virtuellen, menschlichen Form immer realistischer.

### Uncanny Valley

Eine Schwierigkeit bei der Modellierung von menschlichen Kommunikationspartnern im Computer wird mit dem Uncanny Valley ausgedrückt. Die entsprechenden Untersuchungen von M. Mori in 1970 haben untersucht, wie Menschen positiv auf einen menschenähnlichen künstlichen Avatar reagieren. Wenig überraschend ist dabei, dass die Reaktion auf einen gesunden Menschen sehr positiv ausgefallen ist. Offenbar gibt es aber einen Ähnlichkeitsgrad, bei dem die Reaktion sehr negativ ausfällt. Dieser Ähnlichkeitsgrad liegt etwa bei 80 %-90 %. Im Uncanny Valley befinden sich z.B. Leichen oder Zombies. Der Ursprung für die extreme Ablehnung gegen stark menschenähnliche Avatare oder Roboter ist noch ungeklärt. Eine Theorie geht davon aus, dass dies eine natürliche Schutzreaktion ist, durch die der Mensch Kontakt zu kranken Individuen vermeidet. Für bewegte Bilder ist der Effekt des Uncanny Valley deutlich extremer als für stillstehende Bilder.

Uncanny Valley

### Konversationsmodell

Die Kommunikation zwischen Menschen aber auch zwischen Mensch und Computer erfolgt auf verschiedenen Ebenen. Jede dieser Ebenen ist auf eine Weise im Menschen und im Computer repräsentiert. Zwischen den einzelnen Ebenen besteht eine hierarchische

Struktur. Oben stehen sehr intellektuelle Ebenen, die beim Menschen zu einem großen Teil im Bewusstsein stattfinden. Weiter unten befinden sich die Ebenen, die beim Menschen eher unbewusst und beim Computer zu einem Teil direkt in der Hardware stattfinden.:

**Intentionale Ebene** Was soll erreicht werden?

**Pragmatische Ebene** Welche Ziele und Unterziele müssen dazu erreicht werden?

**Semantische Ebene** Welche Gegenstände und Operationen sind dazu nötig?

**Syntaktische Ebene** Wie müssen Äußerungen und Handlungen formuliert werden?

**Lexikalische Ebene** Welche Zeichen und Operationen stehen zur Verfügung?

**Sensomotorische Ebene** Welche motorischen Aktionen müssen durchgeführt werden?

### 5.1.2 Interaktionssicht

Seit 1981 haben sich grafische Benutzerschnittstellen (GUI) immer stärker durchgesetzt. Vorher gab es hauptsächlich Kommandozeileninterfaces (s. Abschnitt 7). Bei diesen Schnittstellen hat es sich durchgesetzt, räumliche Metaphern einzusetzen (z.B. Schreibtisch, Ordner, Archiv, ...). Dadurch wird vom Menschen das System als ein Ort wahrgenommen, an dem sich Gegenstände befinden, mit denen man interagieren kann. So ergibt sich ein Weltmodell, durch das der Computer für den Menschen repräsentiert wird.

Weltmodell

Dadurch, dass die Interaktionssicht so stark durch Metaphern und Symboliken geprägt ist sind hier viel besser deiktische (zeigebasierte) Interaktionsformen möglich, also Interaktion durch eine Maus oder einen Touchscreen. Hier ist im Gegensatz zur Kommunikationssicht weit mehr Wiedererkennen anstelle von Erinnern möglich.

## 5.2 Handlung

Handlung ist umgangssprachlich ein Begriff für beliebige zielgerichtete Bewegungen. Im Kontext der MCI verwenden wir folgende Definition:

### Definition: Handlung

Handlung ist die kleinste Einheit des Verhaltens in Bezug auf ein bewusstes Ziel.

Handlung besteht aus drei Schritten:

Der Ablauf (Abbildung 5.1) findet ständig bei jeder Handlung statt. Der Kreislauf zwischen Durchführung und Beurteilung wird im Allgemeinen so lange durchlaufen, bis die Beurteilung ergeben hat, dass das Ergebnis der Durchführung dem Ziel entspricht.

Wie schon die Kommunikation im vorangegangenen Abschnitt, findet auch die Handlung auf mehreren Ebenen statt:

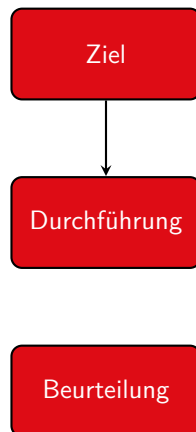


Abbildung 5.1: Handlungssysteme

**Intellektuelle Ebene** Bewusste Regulation

**Sensomotorische Ebene** Automatische Steuerung der Sensorik

**Flexible Handlungsmuster** Kombiniert die beiden oberen Ebenen.

Um die Durchführung einer Handlung zu erleichtern, kann man Werkzeuge verwenden. Die Verwendung von Werkzeugen muss zwar zunächst erlernt und geübt werden, erleichtert danach aber das Erreichen des Handlungsziels. Mit genug Übung verschwindet die Benutzung des Werkzeugs aus dem Bewusstsein und geschieht automatisch.

Aus den drei Ebenen der Handlung leitet sich auch das sogenannte Skill-Rules-Knowledge-Modell ab:

Das Skill-Rules-Knowledge-Modell unterteilt die Handlung in die drei Ebenen

1. Fähigkeitsbasierte Ebene (Skill level)
  - Automatische Handlung, die meist nicht verbalisierbar sind.
  - Reize werden erkannt und ein stereotypisches Handlungsmuster wird ausgeführt.
2. Regelbasierte Ebene (Rule level)
  - Bewusstes Verhalten bei bekannten Situationen, aufgrund von bestimmten Handlungsregeln.
  - Reize werden zu Gedächtnisinhalten zugeordnet, in denen Regeln für die Aufgaben gespeichert sind.
3. Wissensbasierte Ebene (Knowledge level)
  - Verhalten in unbekanntem Situationen, in denen es keine Handlungsregeln gibt.
  - Personen müssen aktiv ein Problem lösen und eine Statusanalyse der Umwelt vornehmen.

### 5.2.1 Human Action Cycle

Das Modell Human Action Cycle wurde entwickelt, um zu analysieren, an welchen Stellen im Handlungsablauf ein Fehler aufgetreten ist, der dazu geführt hat, dass die Handlung nicht korrekt abgeschlossen wurde. Der Human Action Cycle basiert zu großen Teilen auf dem Kommunikationsmodell der Mensch-Computer-Kommunikation. Eine Aktion durchläuft im Human Action Cycle 7 Stufen:

1. Bildung des Ziels.
2. Bildung der Intention und der nötigen Handlungen zum Erreichen des Ziels.
3. Anordnen der Handlungen zu einer Handlungssequenz.
4. Ausführung der Handlungssequenz.
5. Wahrnehmen der Änderungen nach Ausführung der Handlungssequenz.
6. Interpretation des Ergebnisses.
7. Vergleich des Ergebnisses mit dem Ziel aus Schritt 1.

Mithilfe dieses Modells ist es nun möglich, klar zu formulieren, wie ein Fehler entstanden ist. Möchte man beispielsweise ein Auto fernsteuern und fährt aus Versehen rückwärts statt vorwärts, schlägt offenbar in Schritt 7 der Vergleich des Ergebnisses mit dem Ziel fehl. Innerhalb der einzelnen Schritte kann man Fehler auch noch weiter unterteilen. Diese Unterteilungen sind aber meistens sehr aufgabenspezifisch, sodass wir hier nicht weiter darauf eingehen.

#### Aufgabe 15: Human Action Cycle

Welche Prozesse werden im Human Action Cycle durchlaufen, wenn man ein Icon auf dem Desktop löschen möchte?

### 5.3 Fehler

Wir definieren Fehler für uns mithilfe des Human Action Cycle:

#### **Definition: Transformationsdistanz (Kluffen)**

Eine Kluft (*engl. Gulf*) bzw. Transformationsdistanz bezeichnet Schwierigkeiten beim Übergang von einer zur nächsten Handlungsebene im Human Action Cycle.

### Definition: Fehler

Ein Fehler tritt als Konsequenz einer Kluft auf und manifestiert sich in Handlungen, die nicht zielführend sind.

Scheitert im Human Action Cycle in einem Schritt die Ausführung der Aufgabe, liegt ein Fehler vor. Dabei unterscheiden wir

**Gulf of Execution** In diese Kategorie fallen Fehler, die bei der Ausführung einer Ausgabe auftreten. Im Human Action Cycle entspricht dies den Schritten 2-4.

**Gulf of Evaluation** In diese Kategorie fallen Fehler, die bei der Evaluierung des Ergebnisses auftreten. Im Human Action Cycle entspricht dies den Schritten 5-7.

### Beispiel

Angenommen jemand arbeitet mit einem Kommandozeilensystem, hat davor aber nur mit grafischen Benutzerschnittstellen gearbeitet. Möchte dieser Benutzer nun eine Datei löschen, scheitert dies vermutlich am Gulf of Execution und zwar in Schritt 2 im HAC. Der Benutzer kann zwar das Ziel formulieren („Löschen einer Datei“), kennt aber den Befehl nicht und kann somit nicht die Handlungsschritte ableiten.

Ein Beispiel für den Gulf of Evaluation tritt auf, wenn der Benutzer auf ein akustisches Signal des Computers wartet, die Lautsprecher allerdings abgestellt sind. Dann scheitert die Handlung im HAC in Schritt 5, da der Benutzer das Signal des Computers nicht wahrnehmen kann.

Für jeden der 7 Schritte im Human Action Cycle gibt lassen sich die Ursachen für Fehler ableiten. Das ist besonders im Bereich der MCI interessant, da wir durch ein besseres Fehlerverständnis dafür sorgen können, dass dem Benutzer besser geholfen wird, eine Aufgabe zu lösen.

### Aufgabe 16: Transformationsdistanz (Kluffen)

Ordnen sie die folgenden Fehler dem Gulf of Execution bzw. dem Gulf of Evaluation zu und geben Sie den Schritt im Human Action Cycle an, in dem der Fehler auftritt.

1. In welcher Reihenfolge müssen Aktionen erfolgen?
2. Die Icons sind schwer zu treffen.
3. Das Feedback ist nicht lesbar, erkennbar oder unverständlich.
4. Das Ergebnis von Operationen ist unerwartet.

Schritt	Fehler
1. Ziel festlegen	-
2. Absicht formulieren	Zu wenig Kenntnis der Konzepte
3. Aktion planen	Zu wenig Kenntnis der Bedienung
4. Aktion ausführen	Unklare oder umständliche Bedienung
5. Zustand wahrnehmen	Zustand nicht oder nur schwer erkennbar
6. Zustand interpretieren	Darstellung nicht klar interpretierbar
7. Zustandsergebnis festlegen	Zielerreichung kann nicht erkannt werden

Tabelle 5.1: Fehler im Human Action Cycle

### 5.3.1 Fehlerklassifikation

Tabelle 5.1 gibt nur einen groben Überblick über die Schritte im Human Action Cycle, in denen bestimmte Fehler auftreten können. Allgemein klassifizieren wir Fehler nach zwei Kriterien: Dem Zeitpunkt im Model Human Processor und der Fehlerart. Im Model Human Processor können Fehler in allen drei Prozessoren auftreten:

**Perzeptuelle Fehler** treten auf, wenn die Wahrnehmung von Hinweisen fehlerhaft ist. Im HAC entspricht dies hauptsächlich Schritt 5, kann aber auch in Schritt 3 passieren.

**Kognitive Fehler** werden durch Überbeanspruchung kognitiver Ressourcen hervorgerufen (vgl. 4.5)

**Motorische Fehler** entstehen durch mangelnde motorische Fähigkeiten.

Wir unterscheiden folgende Fehlerarten:

**Arbeitsfehler** entstehen durch ein mangelndes Verständnis oder durch ein falsches mentales Modell über ein System. Ein Beispiel für einen Arbeitsfehler ist, wenn ein Benutzer eine Datei löschen möchte, aber stattdessen nur die Verknüpfung zur Datei löscht.

**Flüchtigkeitsfehler (engl. Slip) / Versehen (engl. Lapses)** entstehen durch Unaufmerksamkeit, Ablenkung oder Gewohnheit während der Durchführung (Slip) oder der Verarbeitung (Lapses) einer Aufgabe. Slips und Lapses können wir noch weiter unterteilen:

**Capture Error** Eine vertraute Aktion kapert die eigentlich beabsichtigte Aktion. Voraussetzung dafür ist normalerweise, dass die vertraute Aktion schon stark im prozeduralen Gedächtnis gefestigt ist.

**Description Error** Hier werden Aktionen zwar richtig ausgeführt, allerdings am falschen Zielobjekt. Dies passiert zum Beispiel bei nicht beschrifteten Schaltern für Licht in einem großen Raum.

**Mode Error** Hier wird die korrekte Aktion am richtigen Objekt durchgeführt, allerdings befindet man sich im falschen Modus. Dies ist bei Flug IA-605 (s. Abschnitt 1.3) der Fall gewesen.

**Misslingen/Scheitern (engl. Failure)** entsteht durch die mangelnde Fähigkeit des Benutzers oder durch die Schwierigkeit der Aufgabe, obwohl die Problemlösung bekannt ist.

#### Aufgabe 17: Fehlerklassifikation

Ordnen sie die folgenden Fehler den vorgestellten Fehlerklassifikationen zu:

1. Man öffnet den Internet-Browser, um nach einem Thema zu recherchieren. Stattdessen tippt man „twitter“ in die Adresszeile ein.
2. Man gibt sein Passwort ein, aber die Feststelltaste ist aktiviert.
3. Man nimmt versehentlich statt einem Video ein Foto auf.
4. Man trägt das Passwort im Feld „Benutzername“ ein.
5. Man erhält das falsche Ergebnis beim Multiplizieren zweier Zahlen im Kopf.

## 5.4 Aktion und Motorik

In diesem Abschnitt betrachten den motorischen Prozessor im Model Human Processor (s. Abschnitt 2.3.1) noch etwas genauer.

### 5.4.1 Biologischer Hintergrund

Das motorische System ist der Teil des zentralen Nervensystems, der für die Bewegung zuständig ist. Wie schon beim perzeptuellen System gibt es hier spezielle Nervenbahnen, die Signale vom motorischen Prozessor im Gehirn zu den Muskeln leiten. Die Muskeln führen dann die entsprechende Bewegung aus. Bei den Bewegungen unterscheiden wir zwei Arten:

**Zielmotorik** Die Zielmotorik umfasst bewusste motorische Äußerungen gegenüber der Umwelt (umgangssprachlich: Bewegung). Zielmotorische Bewegungen müssen bewusst geplant werden und werden über das Pyramidenbahnsystem vermittelt. Die Bewegungen werden hauptsächlich durch die distale Muskulatur ausgeführt, die für die Feinbewegungen verantwortlich ist (diese Muskulatur befindet sich an den Endeffektoren wie z.B. den Fingern).

Pyramidenbahnsystem

**Stütz-/Haltemotorik** Diese Motorik findet unbewusst statt und ist für die Aufrechterhaltung der Körperstellung und erlernte Bewegungsautomatismen zuständig. Sie umfasst hauptsächlich proximale Muskulatur (in der Nähe des Rückenmarks) und

wird durch die Propriozeption gesteuert (s. Abschnitt 3.5.1). Die Aktionen der Stützmotorik laufen weitestgehend unbewusst ab und lassen sich kaum beeinflussen.

## 5.4.2 Reafferenzprinzip

Das Reafferenzprinzip verbindet die biologischen Hintergründe der Motorik mit dem Human-Action-Cycle. Beim Reafferenzprinzip gehen wir davon aus, dass bei jeder (ziel)motorischen Aktion neben dem Befehl an den Muskel (Efferenz) eine Efferenzkopie erstellt wird. Die Kopie wird dann zusammen mit dem sensorischen Feedback nach der Bewegung an einen Komparator geleitet, der das Ergebnis mit den Erwartungen vergleicht. Im HAC entspricht der Komparator dem letzten Schritt. Die Efferenzkopie selbst geht nur indirekt aus dem HAC hervor. Sie entspricht dem erwarteten Feedback und ist notwendig, damit der Komparator am Ende das erwartete Ergebnis kennt.

Normalerweise kommt es bei diesem Prozess zu Diskrepanzen. Die Motorik ist normalerweise nicht so exakt, dass genau das erwartete Ergebnis erzielt wird. Daher ist es möglich zusätzlich im Komparator eine Regulation einzuführen. Die Regulation führt dazu, dass eventuelle Folgeaktionen basierend auf dem Ergebnis der aktuellen Aktion anders ausgeführt werden.

### Beispiel: Reafferenzprinzip

Man möchte einen Becher hochheben. Nach dem HAC wird diese Aufgabe zerlegt in zwei Schritte:

1. Becher greifen
2. Becher anheben

Nach dem Reafferenzprinzip erfolgt nach dem ergreifen des Bechers ein Vergleich („umschließt die Hand den Becher?“) und eine Regulation. Die Regulation prüft an dieser Stelle, wie schwer der Becher ist und reguliert die Muskeln entsprechend für den nächsten Schritt. Ist der Becher etwas schwerer als erwartet, gelingt es trotzdem, ihn ohne weitere Probleme anzuheben. Nur wenn das Gewicht des Bechers extrem von den Erwartungen abweicht (viel leichter oder viel schwerer), kann die Regulation dies nicht komplett ausgleichen.

## 5.5 Modelle

Es gibt viele verschiedene Modelle, um Bewegungen zu modellieren. Uns interessieren hier besonders solche Modelle, die für die Mensch-Computer-Interaktion verwendbar sind. Solche Modelle versuchen aufgrund von empirischen Daten vorherzusagen, wie lange es dauert, eine bestimmte Aktion auszuführen. Der Vorteil hierbei ist, dass man so einen numerischen Vergleichswert für zwei Vorgehensweisen (z.B. Maus vs. Joystick) schafft und dadurch besser argumentieren kann, wieso eine Methode besser ist, als eine andere.



### 5.5.1 GOMS

GOMS steht für **G**oals, **O**perators, **M**ethods, **S**election Rules und ist ein spezialisiertes Modell für die menschliche Informationsverarbeitung. Das GOMS-Modell wurde wie auch das Model Human Processor von S. Card, T. Moran und A. Newell entwickelt und zerlegt Aufgaben entsprechend in perzeptuelle, kognitive und motorische Aufgaben.

Der grundsätzliche Ansatz bei GOMS ist, eine Aufgabe in ein bestimmtes Format zu zerlegen. Dieses lässt sich so formulieren:

*„GOMS betrachtet Interaktionsaufgaben als eine durch den Benutzer ausgeführte Selektion aus einer Menge von Methoden, die aus Operationen zusammengesetzt sind, um Ziele zu erfüllen.“*

— S. Card, T. Moran, A. Newell: The Psychology of Human Computer Interaction. Lawrence Erlbaum Associates. ISBN 0-89859-859-1, 1983

Nun gibt es verschiedene Varianten von GOMS. Wir betrachten im Folgenden das CMN-GOMS und das KLM-GOMS.

#### CMN-GOMS

Das CMN-GOMS ist die ursprünglich von **C**ard, **M**oran und **N**ewell entwickelte Interpretation von GOMS. Es interpretiert die einzelnen Elemente im Kontext der MCI so:

**Goals** sind die Ziele des Benutzers. Im HAC entspricht dies Schritt 1. Ein Ziel könnte z.B. das Löschen einer Datei sein.

**Operators** sind die einzelnen Operationen, die nötig sind. Auch dies lässt sich in Schritt 3 im HAC wiederfinden.

**Methods** sind verschiedene Varianten, zum Ziel zu gelangen. Dies entspricht Schritt 2 im HAC. Jede Methode hat andere Operatoren, die dafür nötig sind.

**Selection Rules** sind Regeln, nach denen eine der verfügbaren Methoden ausgewählt wird (falls es mehrere gibt).

Mit dem CMN-GOMS haben wir nun eine formale Methode, um den Ablauf für eine Aktion darzustellen. Normalerweise schreibt man dies im Pseudocode so:

#### Beispiel: CMN-GOMS

Die Aufgabe ist, ein Fenster zu schließen.

```
GOAL: CLOSE-WINDOW
  [select:
    GOAL: CLOSE-METHOD
      MOVE-MOUSE-TO-FILE-MENU
      PULL-DOWN-FILE-MENU
      MOVE-MOUSE-TO-CLOSE-OPTION
```

```

        CLICK-CLOSE-OPTION
    GOAL: CTRL-W-METHOD
        PRESS-CONTROL-W-KEYS
    ]

*Selection rule for GOAL: CLOSE-WINDOW
    If hands are on keyboard, use CTRL-W-METHOD
    else use CLOSE-METHOD

```

Beim CMN-GOMS gehen wir davon aus, dass immer nach den Selection Rules entschieden wird. Je mehr Schritte eine Methode dann hat, desto länger dauert sie (zumindest grob gesehen).

## KLM-GOMS

Das KLM-GOMS (**K**eystroke-**L**evel-**M**odel) ist eine Erweiterung des CMN-GOMS. Im KLM-GOMS muss jede Aktion in eine Abfolge von bestimmten atomaren Aktionen zerlegt werden. Jede dieser atomaren Aktionen hat dabei eine bestimmte Laufzeit, sodass damit am Ende die nötige Gesamtzeit für eine Aufgabe bzw. für jede verfügbare Methode errechnet werden kann.

Die atomaren Aktionen sind in Tabelle 5.2 abgebildet. Die entsprechenden Laufzeiten sind empirisch bestimmt worden und sind daher nur ein Mittelwert.

K	Tastendruck	0,28 s
T(n)	Eintippen einer Sequenz von n Buchstaben auf der Tastatur	$n \times K$
P	Zeigen mit der Maus auf ein Ziel auf dem Display	1,1 s
B	Button-Press oder -Release	0,1 s
BB	Button-Press und -Release	0,2 s
H	Hände zur Tastatur oder zur Maus bewegen	0,4 s
M	Mentaler Akt von Routine-Denken oder Wahrnehmen	1,2 s
W(t)	Warten auf Systemantwortzeit	t

Tabelle 5.2: Atomare Aktionen

Wir können nun das Beispiel aus dem CMN-GOMS mit dem KLM-GOMS erweitern:

### Beispiel: KLM-GOMS

CLOSE-METHOD		CTRL-W-METHOD	
P (to Menu)	1,1 s	M	1,2 s
B (LEFT down)	0,1 s	K (CTRL & W)	0,28 s
M	1,2 s		
P (to Close)	1,1 s		
B (LEFT down)	0,1 s		
<b>Total:</b>	<b>3,6 s</b>	<b>Total:</b>	<b>1,48 s</b>

Durch das KLM-GOMS können wir damit für sehr viele Methoden sehr einfach abschätzen, wie lange das Ausführen der Methoden braucht. Dadurch können verschiedene Methoden einfach miteinander verglichen werden.

Mit dem KLM-GOMS können aber auch Mensch-Computer-Interaktionen qualitativ bewertet werden:

- Viele Ms sind Hinweise auf Stockungen im Arbeitsablauf (der Benutzer muss häufig nachdenken).
- Viele Hs deuten auf einen zu häufigen Wechsel zwischen Maus und Tastatur hin.
- Viele Ps anstelle von wenigen Ks deuten auf eine langsame Bedienung hin.

### 5.5.2 Fitts' Law

Im GOMS-Modell werden Bewegungen im Mittel mit einer Zeit von  $P = 1,1$  s angegeben. Im Mittel ist dies auch ein sehr guter Wert, allerdings kann man sich schnell klar machen, dass es offenbar deutlich länger dauert, eine große Strecke mit der Maus zurückzulegen als eine sehr kurze. Andererseits kann es auch sehr lange dauern, ein kleines Ziel zu treffen, selbst wenn sich dieses sehr nahe am Mauszeiger befindet.

Diese Beobachtungen können wir mithilfe der Aspekte **Geschwindigkeit** (des Mauszeigers) und **Genauigkeit** formalisieren. Allgemein gilt, dass die Geschwindigkeit mit zunehmender Genauigkeit abnimmt. Das gilt auch für den Mauszeiger am Computer. In der MCI geht es darum, das richtige Verhältnis zwischen Geschwindigkeit und Genauigkeit zu finden, damit der Benutzer Aktionen möglichst schnell durchführen kann.

#### Definition: Fitts' Law

Die Zeit  $T$  (in Sekunden), die benötigt wird, um mit der Maus ein Ziel der Größe  $S$  in der Distanz  $D$  zu treffen, berechnet sich durch:

$$T = a + b \cdot \log_2 \left( \frac{D}{S} + 1 \right)$$

$a$  und  $b$  sind dabei empirisch bestimmte Werte.  $a$  wird auch **intercept** genannt

und ist die Mindestantwortzeit für eine Bewegung.  $b$  wird auch **slope** genannt und ist der Faktor, mit dem die Aufgabe bei zunehmender Distanz bzw. abnehmender Größe schwieriger wird.

Die Variablen  $a$  und  $b$  sind von Person zu Person, aber auch in unterschiedlichen Situationen sehr unterschiedlich. Faktoren, die auf  $a$  und  $b$  Einfluss haben, sind zum Beispiel die Mausgeschwindigkeit, die Mausform, die Umgebungshelligkeit oder die Bildschirmhelligkeit. Da diese Variablen somit nur für eine sehr bestimmte Situation viel Aussagekraft haben, verwendet man statt der kompletten Formel von Fitts auch of den Index of Difficulty (ID) oder den Durchsatz (Index of Performance, IP). Diese sind folgendermaßen definiert:

Index of Difficulty

$$ID = \log_2 \left( \frac{D}{S} + 1 \right)$$
$$IP = \frac{ID}{T}$$

Der Index of Performance hat für reelle Szenarien die größte Aussagekraft, da mit diesem angegeben wird, wie viel schneller als  $ID$  die Aufgabe gelöst wurde.

### Effektiver Durchsatz

Mit dem Durchsatz haben wir zwar einen Wert, der sich für den Vergleich reeller Experimente eignet, allerdings wird dabei außer Acht gelassen, dass normalerweise weder die Distanz noch die Größe eines Objektes vollständig ausgenutzt wird. Führt man beispielsweise ein Experiment nach Fitts durch, bei dem die Zielobjekte sehr groß sind, werden die meisten Menschen die Maus gar nicht bis zur Objektmitte bewegen, sondern nur bis zum am nächsten gelegenen Objektrand. Dadurch ist die effektive Distanz geringer als die tatsächlich gemessene Distanz. Ähnlich verhält es sich bei der Größe von Objekten. Liegt ein Objekt schief oder hat eine ungewöhnliche Form, entspricht die effektive Größe nicht der tatsächlich gemessenen Größe.

Damit erhalten wir neue Formeln für den effektiven ID und den effektiven IP (effektiver Durchsatz):

$$ID_e = \log_2 \left( \frac{D_e}{S_2} + 1 \right)$$
$$IP_e = \frac{ID_e}{T}$$

### Grenzen von Fitts' Law

Fitts' Law gilt nur für Bewegungen, die aus einer ballistischen Phase und einer Korrekturphase bestehen. In der ballistischen Phase findet eine schnelle, ungenaue Bewegung statt. Erst in der Korrekturphase wird die Bewegung so korrigiert, dass die nötige Präzision vorhanden ist. Zeichnet man einen Graphen, der die Geschwindigkeit der Bewegung

gegen die Zeit aufträgt, erkennt man die beiden Phasen sehr gut daran, dass die Geschwindigkeitswerte in der ballistischen Phase sehr hoch sind, in der Korrekturphase dagegen sehr klein. Zwischen den beiden Phasen ist die Geschwindigkeit fast Null.

#### Aufgabe 18: Fitts' Law

### 5.5.3 Steering Law

Bei grafischen Benutzerschnittstellen ist es sehr populär, Popup Menüs zu verwenden. Diese haben auch auf die Mausbewegung Einfluss, da hier Fitts' Law nur bedingt gilt. Insbesondere wenn man ein Untermenü öffnen will, muss man bei den meisten Betriebssystemen den Mauszeiger in einem schmalen Pfad navigieren, damit das Untermenü geöffnet bleibt. Solche Bewegungen lassen sich durch das Steering Law beschreiben. In seiner allgemeinsten Form lautet es so:

#### Definition: Steering Law

Die Zeit  $T$ , die benötigt wird, um den Mauscursor durch einen zweidimensionalen Pfad zu bewegen, berechnet sich durch

$$T = a + b \cdot \int_S \frac{1}{W(s)} ds$$

dabei sind wie bei Fitts' Law  $a$  und  $b$  empirisch bestimmte Werte.  $W(s)$  ist die Breite des Pfades (orthogonal zur Bewegungsrichtung der Maus) an der Stelle  $s$ .

In der Mensch-Computer-Interaktion betrachten wir normalerweise rechteckige Pfade. Das hat den Grund, dass die meisten Elemente auf Computerbildschirmen rechteckig sind (zum Beispiel Popup Menüs). Für solche Pfade vereinfacht sich die Formel zu

$$T = a + b \cdot \frac{W}{H}$$

Mit der Weite des Pfades  $W$  (entlang der Bewegungsrichtung) und der Höhe  $H$  (rechtwinklig zur Bewegungsrichtung). Für die Werte  $a$  und  $b$  gilt wie bei Fitts' Law, dass diese sehr stark von der aktuellen Situation abhängen.

#### Aufgabe 19: Steering Law

Aus Fitts' Law und dem Steering Law ergeben sich direkt Konsequenzen für die Mensch-Computer-Interaktion:

- Ziele sollten nicht zu klein sein. Die Ziele müssen erkannt und gefunden werden.

- Bei fortlaufenden Aktionen sollten Ziele nahe beieinander sein.
- Häufig gesuchte Ziele sollten möglichst immer an der gleichen Stelle sein.
- Möglichst wenig weit entfernte Ziele.

#### Aufgabe 20: Schlussfolgerungen

Begründen Sie die obigen Konsequenzen für die Mensch-Computer-Interaktion mit Hilfe der in diesem Kapitel behandelten Modelle.

### 5.5.4 CLC-Model

Das CLC-Model verwendet das KLM-GOMS sowie Fitts' Law und das Steering Law sowie weitere Erkenntnisse, um die Dauer für die Ausführung von komplexeren Bewegungen zu berechnen. Der Name steht für **C**urves, **L**ine segments and **C**orners. Hier werden dann Bewegungen in Bewegungen aus fünf verschiedene Kategorien zerlegt, die in Abbildung 5.2 dargestellt sind. Für die Berechnung unterscheidet man dann nach Kategorie, da z.B. eine gradlinige Bewegung im Allgemeinen schneller auszuführen ist, als eine Kreisbewegung.

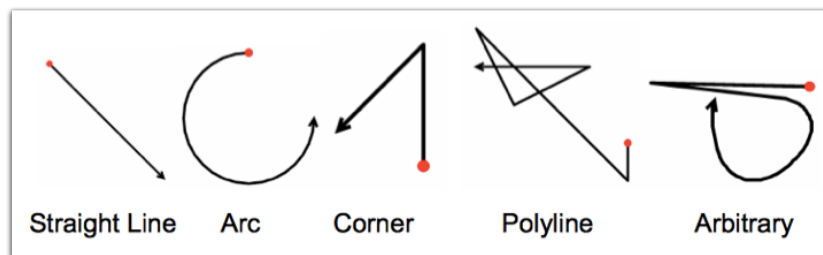


Abbildung 5.2: CLC-Model

## 5.6 Kontrolle-Display-Relation

Kontrolle-Display-Relation ist ein Begriff aus der Mensch-Computer-Interaktion, mit dem das Verhalten eines Eingabegerätes beschrieben werden kann. Der Grundgedanke dabei ist, dass jede Bewegung eines Eingabegerätes mit einer virtuellen Bewegung auf dem Bildschirm korrespondiert und entsprechend übersetzt wird. Die Kontrolle-Display-Relation ist die entsprechende Transferfunktion zwischen physischem Gerät (z.B. Maus) und virtueller Repräsentation (z.B. Mauszeiger).

#### Definition: Display

Ein Display ist in der Mensch-Computer-Interaktion ein Ausgabegerät. Obwohl man

umgangssprachlich damit meistens nur den Bildschirm bezeichnet, umfasst dies auch Lautsprecher, haptische Geräte sowie beliebige andere Wege über die der Computer Feedback an den Menschen sendet.

### 5.6.1 Räumliche Relation

Die räumliche Relation ist die häufigste Art von Kontrolle-Display-Relation. Als Display verwenden wir hier meistens den Bildschirm, allerdings kann man genauso die Lautsprecher oder jedes andere Ausgabegerät verwenden. Wir stellen uns das Display als ein Ausgabegerät mit verschiedenen räumlichen Attributen vor. Beim Bildschirm bieten sich meistens die drei Translationsachsen  $x$ ,  $y$  und  $z$  sowie drei Rotationsachsen  $\theta_x$ ,  $\theta_y$  und  $\theta_z$  an. Wir verwenden hier ein linkshändriges Koordinatensystem, bei dem die  $x$ -Achse auf dem Bildschirm einer links-rechts-Bewegung entspricht, die  $y$ -Achse einer hoch-runter-Bewegung und die  $z$ -Achse einer reinzoomen/rauszoomen-Bewegung. Besonders an der  $z$ -Achse kann man sehen, dass diese Achsen keine physischen Attribute des Ausgabegerätes sein müssen. Es reicht, wenn das Ausgabegerät eine Achse simulieren kann.

Räumliche Kontrolle-Display-Relationen werden zum Beispiel bei der Maus angewendet. Die physische Maus kann in drei Axen bewegt werden (links/rechts, vorne/hinten, hoch runter). Wie wir alle wissen wird eine links/rechts-Bewegung der Maus in eine links/rechts-Bewegung auf dem Bildschirm, eine vorne/hinten-Bewegung aber in eine oben/unten-Bewegung auf dem Bildschirm übersetzt. Um dies auszudrücken, gibt es eine Notation für Kontrolle-Display-Relationen (Abbildung 5.3). Für die Maus verwenden wir das gleiche Koordinatensystem wie für den Bildschirm (d.h. ein typischer Schreibtisch befindet sich in der  $x$ - $z$ -Ebene).

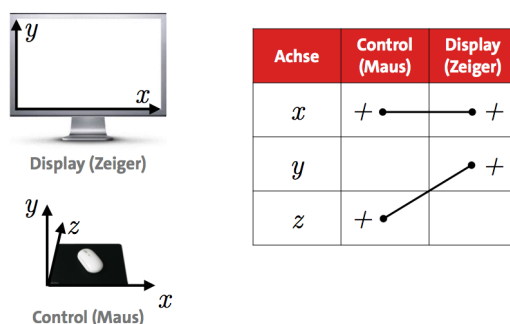


Abbildung 5.3: C/D-Relation (Maus)

Die Notation kann weitergeführt werden, um auch virtuelle Umgebungen wie zum Beispiel Scroll-Panes zu unterstützen. Hier spricht man von Soft Control, da sie nicht mit physischen Eingabegeräten korrespondieren. In Abbildung 5.4 ist die Mausbewegung in Kombination mit Scrollpanes dargestellt.

Soft Control

Für andere Displays als den Bildschirm lassen sich ähnliche C/D-Tabellen erstellen. Zum Beispiel kann die Lautstärke über einen Slider oder ein Drehrad eingestellt werden. Die entsprechenden C/D-Relationen sind in Abbildung 5.5 dargestellt.

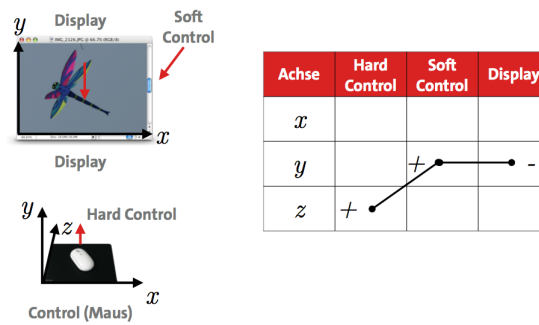


Abbildung 5.4: C/D-Relation (Scrollpane)

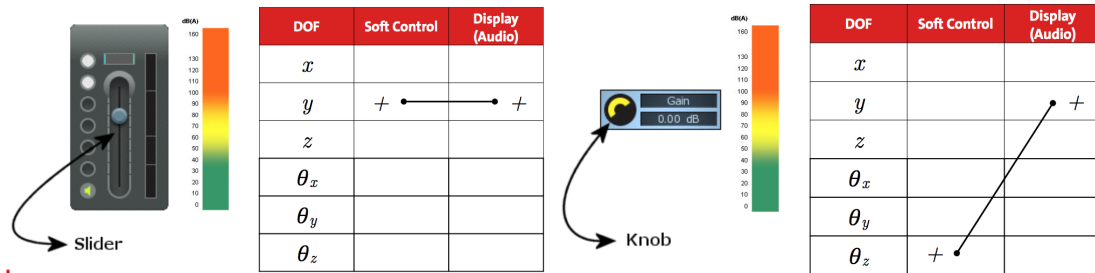


Abbildung 5.5: C/D-Relation (Audio)

## Aufgabe 21: C/D-Relation

### 5.6.2 C/D-Sensitivität

Mit dem oben beschriebenen Modell für räumliche Relationen kann man die Kontrolle-Display-Relation nur qualitativ beschreiben. Häufig ist es aber wünschenswert auch eine quantitative Beschreibung der Relation zu haben. Dies kann man über die C/D-Sensitivität machen.

#### Definition: C/D-Sensitivität

Die C/D-Sensitivität beschreibt die Sensitivität der Kontrolle-Display-Relation. Sie wird über C/D-Gains definiert. C/D-Gains sind multiplikative Konstanten, die mit der Kontroll-Bewegung multipliziert werden, bevor die Bewegung auf dem Display ausgeführt wird.



### Beispiel: C/D-Gains

Ein C/D-Gain von 2 bei der Maus bedeutet, dass eine physische Bewegung von 1 cm in eine Bewegung von 2 cm auf dem Bildschirm übersetzt wird.

Bei der Bestimmung der optimalen C/D-Sensitivität für ein Kontroll-Gerät gilt es immer den **Geschwindigkeits-Genauigkeit-Konflikt** zu minimieren. Diesen haben wir schon bei Fitts' Law kennen gelernt (s. Abschnitt 5.5.2) kennen gelernt. Auch bei anderen Kontroll-Geräten als der Maus kommt es normalerweise dazu, dass eine Bewegung entweder sehr schnell oder sehr präzise durchgeführt werden kann. Im Allgemeinen sind große C/D-Gains für sehr grobe Aufgaben sehr gut geeignet, aber kleine C/D-Gains für sehr präzise Aufgaben notwendig. Jeder Mensch muss letztlich für sich selbst die optimalen Gains finden, bei denen der Geschwindigkeits-Genauigkeit-Konflikt möglichst minimal wird.

### 5.6.3 Three-State-Model

Das Three-State-Model ist eine Möglichkeit, Zustände von Eingabegeräten und Übergänge dazwischen zu beschreiben. Das Modell verwendet eine an Automaten angelehnte Notation. Das Three-State-Model für eine Maus ist in Abbildung 5.6 abgebildet.

Abbildung 5.6: Three-State-Model

Ein Problem am Three-State-Model ist, dass die drei erlaubten Zustände nicht immer so einfach zu definieren sind, wie bei der Maus. Zum Beispiel kann bei einer Eingabe über einen Touchscreen nur schwer zwischen dem Tracking- und Dragging-Status unterschieden werden. In der Konsequenz wird das Modell teilweise erweitert und es werden neue Zustände eingeführt, um ein Eingabegerät besser zu beschreiben.

**Positionierung** Das Problem, das beim Vergleich zwischen Maus und Touch-Eingabe im Three-State-Model auftritt, kann auf die Art der Positionierung zurückgeführt werden. Wir unterscheiden zwei Arten:

**Absolute Positionierung** bildet die Position der Kontrolleingabe absolut auf der Displayausgabe aus. Dies ist beim Touchscreen der Fall, bei dem man die Ausgabe des Displays genau an der Stelle erfolgt, an der man getippt hat.

**Relative Positionierung** bildet die Kontrollbewegungen relativ auf das Display ab. Dies ist bei der Maus oder einem Trackpad der Fall, bei denen eine Bewegung den Mauszeiger relativ zur vorhergegangenen Position bewegt.

# 6 Historie der MCI

## 6.1 Generationen der Informatik

Informatik im heutigen Sinne wurde erst in der Mitte des 20. Jahrhunderts begründet. Allerdings gab es schon vorher Gedanken und theoretische Betrachtungen im Bereich automatisiertes Rechnen. Eine der ersten solcher Konzepte entstand im 17. Jahrhundert durch Leibniz. Dieser entwickelte 1673 eine Mechanische Rechenmaschine, die addieren, subtrahieren und multiplizieren konnte. Leibniz' Rechenmaschine ist eine der ersten Maschinen, bei denen klar das Prinzip *Eingabe, Verarbeitung, Ausgabe* erkennbar ist. Damit hat er einen wichtigen Grundstein für das Entstehen der Informatik gelegt. Leibniz machte sich auch schon Gedanken über die Zweckmäßigkeit des Binärsystems für automatisiertes Rechnen und hat letztlich ein vollständiges Kalkül für das Rechnen im Binärsystem erstellt.

1804 entwickelte J.M. Jacquard eine Methode, wie man Webstühle durch Lochkarten *programmieren* konnte. Über die Lochkarten konnte man einstellen, welche Muster durch den Webstuhl erzeugt werden sollten. Auch hier ist das EVA-Prinzip erkennbar: Die Lochkarten als Eingabe, das Weben ist die Verarbeitung und das fertige Produkt die Ausgabe. Diese Erfindung von Jacquard stellt ebenfalls einen Meilenstein auf dem Weg zur heutigen Informatik dar, da die eines der ersten Beispiele in der Geschichte ist, bei denen eine Maschine abhängig von der Eingabe unterschiedliche Bewegungsabläufe ausführen kann.

### 6.1.1 Hardware

#### Relais

1941 stellte Konrad Zuse den Z3 in Berlin vor. Der Z3 war der erste programmgesteuerte Binärrechner mit einer Speicher- und Zentralrecheneinheit (CPU). Programmiert wurde er entweder über ein fest verbautes Programm oder über Lochkarten. Intern verwendete er Relais (s. Abbildung 6.1). Ein Relais besteht aus drei Kernkomponenten: Einer Spule, einem Anker und einem Kontaktschalter. Wird an der Spule ein Strom angelegt, wird der Anker elektromagnetisch angezogen und sorgt über einen Hebel dafür, dass der Kontaktschalter geschlossen wird. Relais arbeiten somit grundsätzlich auf mechanische Weise, um zwei Zustände 0 und 1 (offen und geschlossen) zu repräsentieren. Sie sind etwa 10 cm groß. Abgesehen von der Größe ist auch die grundsätzliche Funktionsweise von Relais relativ unsicher, da das magnetische Feld einer Spule den Anker eines benachbarten Relais anziehen kann, sodass Fehler im System entstehen.

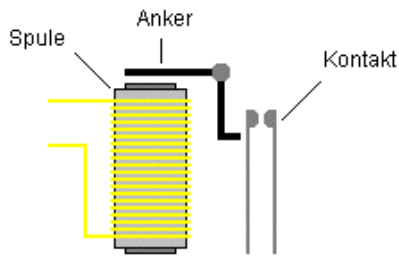


Abbildung 6.1: Relais

**Bugs** Mit dem Begriff Bug wird ein Fehler in einem Computerprogramm bezeichnet. Der Begriff geht ursprünglich auf die Anfänge der Telekommunikation zurück. Bei einer instabilen Telefonleitung hörten die Gesprächspartner ein Kratzen in der Leitung, das an das Krabbeln eines Käfers erinnert. Der Mythos sagt, dass Grace Hopper im Jahr 1945 einen echten Käfer (*engl. Bug*) in einem Relais fand, und dadurch den Begriff für Fehler in Computerprogrammen prägte.

Bug

### Vakuurröhren

Vakuurröhren wurden parallel zu Relais als alternative Technologie entwickelt. Im Jahr 1942 wurde mit ENIAC der erste rein elektronische Universalrechner vorgestellt. Dieser arbeitete mit ca. 18 000 Vakuurröhren (s. Abbildung 6.2). Durch das Anlegen von Strom an die Pins einer Vakuurröhre kann auch dort der Stromfluss gestartet bzw. unterbrochen werden. Im Gegensatz zu Relais haben Vakuurröhren keine beweglichen Teile, arbeiten aber weiterhin mit magnetischen Feldern. Dadurch konnten Vakuurröhren deutlich schneller zwischen 0 und 1 hin- und herschalten. Außerdem waren sie mit 5 cm-8 cm etwas kleiner.

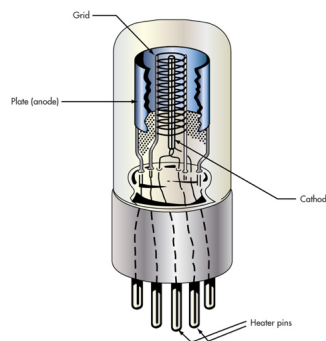


Abbildung 6.2: Vakuurröhre

## Von-Neumann-Architektur

Bei der Entwicklung des ENIAC waren nicht nur die Vakuumröhren eine revolutionäre Technologie, sondern auch die gesamte Architektur des Computers. Diese wurde von J. von Neumann entwickelt und teilt den Rechner in fünf Komponenten auf:

**Arithmetic Logic Unit** Das Rechenwerk führt die tatsächlichen Berechnungen durch.

**Control Unit** Das Steuerwerk bestimmt welche Berechnungen in welcher Reihenfolge durchgeführt werden. Das Steuerwerk kann bedingte Berechnungen auf Grundlage von vorherigen Ergebnissen durchführen.

**Bus-System** Das Bus-System verbindet alle Komponenten im Computer miteinander.

**I/O Unit** Dies ist die Benutzerschnittstelle zum Menschen, über die die Ein- und Ausgabe erfolgt.

**Memory** Im Speicher können Ergebnisse von Berechnungen als Zwischenergebnis abgelegt werden.

Die Von-Neumann-Architektur wird auch heute noch in modifizierter Form in Computern angewendet. Da moderne Computer deutlich schneller sind, ist die Architektur besonders im Bereich des Speichers anders, um zwischen Festplatten, RAM und Caches unterscheiden zu können.

Besonders neuartig an der Von-Neumann-Architektur war, dass hier erstmalig der Programmcode gemeinsam mit den Daten in demselben Speicher gelagert sind. Das Steuerwerk lädt die einzelnen Befehle und Daten aus dem Speicher und gibt diese dann ans Rechenwerk weiter. Dies war historisch gesehen der erste Schritt zu moderner Software und Allzweckcomputern, wie wir sie heute kennen.

## Transistoren

Transistoren sind die Technologie, mit der Computer heute gebaut werden. Durch das Anlegen einer Spannung an die Basis des Transistors, kann der Stromfluss vom Emitter zum Collector unterbrochen werden (s. Abbildung 6.3). Auf diese Weise können wieder die Zustände 0 und 1 repräsentiert werden. Transistoren sind deutlich schneller als Vakuumröhren (etwa um den Faktor 100) und auch deutlich kleiner. Die ersten Transistoren waren wenige cm groß, die aktuelle Generation von Intel-Prozessoren verwendet Transistoren einer Größe von 14 nm. Durch die Transistortechnologie ist es möglich geworden, integrierte Schaltkreise mit mehreren Milliarden Transistoren zu bauen, die die enormen Leistungen moderner Computer ermöglichen.

**Moore's Law** Nachdem sich Transistoren als dominierende Technologie für Computer durchgesetzt haben, hörte die Geschwindigkeitssteigerung nicht auf. Dies führte dazu, dass Gordon E. Moore 1965 ein Gesetz über die Anzahl der Transistoren auf einem Chip formulierte:

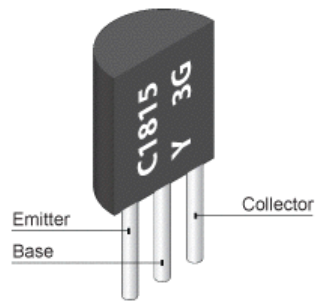


Abbildung 6.3: Transistor

**Definition: Moore's Law**

Alle 1,5 Jahre verdoppelt sich die Anzahl von Transistoren, die in einem integrierten Schaltkreis verbaut werden können.

Die Vorhersage von Moore hat sich bis heute bewahrheitet. Allerdings ist klar, dass das Gesetz auch Grenzen hat, da ein Transistor nicht unendlich weit verkleinert werden kann. Und tatsächlich sind wir mit unserer heutigen Technologie schon sehr nahe an dieser Grenze. Weitere Leistungssteigerungen konnten dann noch durch die Erfindung von Multi-Core Prozessoren erreicht werden, allerdings umfasst Moore's Law eigentlich nur einen einzelnen integrierten Schaltkreis.

Betrachtet man aber die Entwicklung der Computertechnologie unabhängig von der Transistor-Technologie, stellt man fest, dass Moore's Law schon seit vielen Jahrzehnten gilt. Die Steigerung von Relais über Vakuumröhren bis hin zu Transistoren entspricht ebenso einem exponentiellen Wachstum. Dies legt auch die Vermutung nahe, dass sich die Technologie weiter entwickeln wird und Moore's Law auch in Zukunft seine Gültigkeit behält. Die entsprechenden Technologien werden aber vermutlich nicht mehr auf Transistoren basieren (an Alternativen wie z.B. Quantencomputern wird aktuell schon geforscht).

**Aufgabe 22: Moore's Law**

Aktuell befinden sich in einem normalen Computer etwa 5 Milliarden Transistoren. Das menschliche Gehirn hat etwa 100 Milliarden Nervenzellen. Berechnen Sie unter der Annahme, dass Moore's Law weiterhin gültig bleibt, wann ein normaler Computer mehr Transistoren hat, als sich im menschlichen Gehirn Nervenzellen befinden. Nehmen Sie an, dass sich die Anzahl der Nervenzellen im menschlichen Gehirn nicht verändert.

### 6.1.2 Software

Bis 1945 die Von-Neumann-Architektur erfunden wurde, wurden alle Computer entweder über verdrahtete Hardware oder über Lochkarten programmiert. Dadurch hatte auch jeder Computer mindestens zwei Arten von Benutzerschnittstellen: Eine für den Programmierer und eine für den Nutzer. Durch das neuartige Konzept, dass sowohl Programme als auch Daten in demselben Speicher liegen können, wurde es nun möglich, Computer im heutigen Sinne zu programmieren. Somit wurde die zweite Generation der Informatik begründet, die Software-Generation.

In der Software-Generation beschäftigte man sich unter Anderem mit theoretischen Konzepten, wie der Einsatz von Computern im Alltag aussehen könnte. Im Jahr 1945 brauchten Computer noch zig Quadratmeter Fläche und waren somit nicht alltagstauglich. Erst im Laufe der zweiten Generation der Informatik konnte sich die Idee eines privaten Computers (PC) durchsetzen.

#### MEMEX

Der MEMEX (**M**emory **E**xtender) ist ein konzeptionelles Gerät, welches 1945 von V. Bush entwickelt wurde. Es hat die folgenden Funktionen:

- Speicherung sämtlicher Bücher, Bilder und Dokumente
- Gezielter und flexibler Zugriff auf alle Daten
- Anlegen von Verbindungen (Links) zwischen Daten

Auch wenn der MEMEX nur ein konzeptionelles Gerät war, welches in der ursprünglichen Form nie umgesetzt wurde, lässt sich hieraus bereits ein Paradigma der Mensch-Computer-Interaktion ableiten:

- Maschinelle Unterstützung „natürlichen“ Denkens und Verstehens
- Maschinelle Erweiterung zur Unterstützung von Gedächtnis und spontanen Assoziationen
- Organisation und Manipulation analog zur assoziativen Informationsverarbeitung beim Menschen.

Die Online-Enzyklopädie Wikipedia arbeitet genau nach diesem Paradigma und kann somit als moderne Umsetzung des MEMEX betrachtet werden.

#### Man-Computer Symbiosis

*„The hope is that [...] human brains and computing machines will be coupled together very tightly and that the resulting partnership will think as no human brain has ever thought and process data in a way not approached by the information-handling machines we know today.“*

— J.C.R. Licklider: Man-Computer Symbiosis, 1960

Das obige Zitat macht qualitativ sehr deutlich, was mit Man-Computer Symbiosis gemeint ist. Im Detail unterteilte Licklider die Idee in kurzfristige, mittelfristige und langfristige Ziele. Kurzfristige Ziele waren:

**Time-Sharing** Wie können möglichst viele Leute gleichzeitig an einem System arbeiten, ohne dass es zu Konflikten kommt?

**Ein-/Ausgabe** Speicherung und Abruf von großen Datenmengen.

**Interaktive *Real-Time* Systeme** Ergebnisse, die in Echtzeit an den Benutzer geliefert werden.

Mittelfristige Ziele:

- Kooperation bei der Entwicklung großer Softwareprojekte
- Kombination von Sprach- und Schrifterkennung

Langfristige Ziele:

- Verstehen natürlicher Sprache
- Heuristische Programmierung und künstliche Intelligenz

### 6.1.3 Mensch-Computer-Interaktion

**Whirlwind** Das Whirlwind-System wurde 1951 als Flugsimulator am MIT entwickelt. Es war der erste Computer mit einem Bildschirm. Der Bildschirm verwendete eine Kathodenstrahlröhre zum Anzeigen zweifarbiger Grafiken. Das System unterstützte auch einen Lichtgriffel als Eingabegerät, mit dem man auf dem Bildschirm Zeichnen und Objekte verschieben konnte.

**SAGE** Der Whirlwind wurde dann 1958 von der U.S. Air Force zum SAGE Air Defense System (**S**emi-**A**utomatic **G**round **E**nvironment) weiterentwickelt und wurde bis 1983 verwendet. Damit hält das SAGE-System den Weltrekord des am längsten aktiv verwendeten Computersystems.

**Sketchpad** 1962 wurde am MIT ein Lincoln-TX-2 Rechner zum Sketchpad-System umgebaut. Auf dem Sketchpad-System konnten sogar Benutzer, die sich nicht detailliert mit Computern auskannten, mit einem Lichtgriffel Bilder zeichnen und so mit dem Computer interagieren. Das MIT Sketchpad setzte bereits sehr früh objektorientierte Ansätze um und stellt einen wesentlichen Einfluss auf die Entwicklung der Mensch-Computer-Interaktion dar.

**oN-Line System** 1968 wurde am Stanford Research Institute das NLS (oN-Line System) entwickelt, ein interaktives System, welches die Bearbeitung eines Textdokumentes durch zwei Personen gleichzeitig an zwei unterschiedlichen Konsolen ermöglicht hat. Dieses System setzte verschiedene neuartige Technologien um wie zum Beispiel mehrere Fenster, On-Screen Tele-Conferencing, die Maus, eine Tastatur, . . . Im Jahr 1968 stellte Douglas Engelbart das System und all die Technologien in der „Mother of all Demos“ vor.

Mother of all Demos

## 6.2 Klassische Eingabegeräte

Genau wie die Technologie, die zum Bau von Computern verwendet wurde, entwickeln sich auch die Schnittstellen zwischen Mensch und Computer ständig weiter. Bis hin zu modernen Lasermäusen und Touchscreens gab es viele Varianten von Eingabegeräten, von denen sich nur wenige durchgesetzt haben. Wir beschränken uns hier auf die Entwicklung der klassischen Eingabegeräte Maus und Tastatur. Neben diesen beiden Geräten gab es noch viele weitere Forschungen zu Eingabemethoden, die sich jedoch nicht in der breiten Masse durchsetzen konnten.

### 6.2.1 Maus

Die ersten Entwicklungen, die bis zur heutigen Maus führen, ergaben sich am Anfang der 1950er Jahre, als es durch die Von-Neumann-Architektur möglich wurde, Computer zu programmieren, ohne fest verdrahtete Hardware umstecken zu müssen. Zu dieser Zeit stand man vor dem Problem, eine Benutzerschnittstelle zu entwickeln, über die man mit dem Computer kommunizieren konnte, um diesen zu programmieren. Ein Ansatz dafür war der DATAR.

DATAR

**DATAR** Der DATAR wurde ursprünglich von kanadischen Militär entwickelt. Er besteht aus einer frei beweglichen Kugel, an der über zwei Potentiometer die Drehrichtung und -geschwindigkeit gemessen wird. In mechanischen Mäusen wird ein sehr ähnliches Setup mit einer Rollkugel verwendet. Das DATAR-Projekt war allerdings um einiges größer als eine heutige Computermaus. Als Kugel wurde eine kanadische Bowlingkugel verwendet, die händisch in die richtige Richtung gedreht werden musste.

**Grafacon** Das Grafacon war ein etwas kleineres Gerät als der DATAR und konnte bereits auf dem Schreibtisch stehen. Er verwendete ein lineares Potentiometer für Benutzereingaben und als eine art physischer Slider gesehen werden.

**Light-Pen** Der Light-Pen war das Eingabewerkzeug, das 1962 zusammen mit dem Sketchpad-System entwickelt wurde. Die Technologie des Light-Pen basiert auf der Funktionsweise von Elektronenstrahlmonitoren: Bei diesen Monitoren wird ein Elektronenstrahl so gelenkt, dass er mit 120 Hz jeden Punkt des Bildschirms erreicht. Bei einem schwarz-weiß-Display wird der Lenkmechanismus des Elektronenstrahls dauerhaft



betrieben, allerdings wird der Elektronenstrahl selbst nur dann bis zur Anzeigefläche weitergeleitet, wenn an der aktuellen Position ein Pixel aktiviert werden soll. Der Light-Pen konnte nun durch Fototransistoren den Elektronenstrahl erkennen. So konnte berechnet werden, wo er gerade hinzeigte.

**Die Maus** 1963 stellte Douglas Engelbart in der „Mother of all Demos“ eine erste Version der Maus vor. In dieser ersten Version bestand das Gerät aus einem Holzkasten und zwei orthogonal zueinander angeordneten Rädern (also ohne Rollkugel). Obwohl sich die Bedienung der Maus durch die orthogonalen Räder als relativ kompliziert herausstellte, waren doch viele Menschen sehr beeindruckt von der intuitiven Art der Bewegung.

Parallel zu Douglas Engelbart wurde das Prinzip der Maus auch bei der Firma Telefunken erforscht. Hier wurden allerdings keine Räder, sondern ähnlich wie beim DATAR eine Rollkugel verwendet. Diese Variante der Maus konnte man damit deutlich einfacher auf dem Tisch hin und her bewegen. Bei Telefunken entschied man sich allerdings dafür, die Erfindung nicht zu patentieren, da man sich keinen großen Nutzen von der „Rollkugel“ versprach. Das Prinzip der Rollkugel hat sich allerdings für Mäuse durchgesetzt.

**Der Joystick** Als eine Alternative zur Maus wurde 1967 der Joystick vorgestellt. Dieser hatte schon damals große Ähnlichkeit mit heutigen Joysticks. Die Interaktion mit dem Computer über den Joystick erfolgte entweder über relative oder absolute Positionierung. Bei der relativen Positionierung bewegte sich die Maus so lange in eine Richtung, bis man den Joystick wieder in die Ausgangsstellung brachte. Bei der absoluten Positionierung bewegte sich die Maus sofort an einen Punkt, der aus dem aktuellen Winkel des Joysticks berechnet wurde.

**Knee brace** Ebenfalls als Alternative zu Maus und Joystick wurde 1965 die Knieklammer entwickelt. Dieses Eingabegerät befand sich an der Unterseite des Tisches und wurde mit dem Knie bedient. Man konnte mit dem Knie einen hebel hin und her bewegen, was auf dem Bildschirm wiederum in Bewegungen des Cursors übersetzt wurde.

In den 1960er Jahren wurden Computer hauptsächlich für Textverarbeitung verwendet. Daher wurde 1967 eine Untersuchung durchgeführt, bei der verschiedene Eingabegeräte für die Textauswahl verglichen wurden (Maus, Knee brace, Joystick, Grafacon und Light Pen). Dabei hat man einerseits die Zeit gemessen, die für das Auswählen von Wörtern oder einzelnen Zeichen gebraucht wurde und andererseits die Fehlerrate der Benutzer. Diese Experimente haben klar gezeigt, dass die Maus den anderen Eingabegeräten in allen Punkten überlegen war. Mit der Maus konnten die Benutzer Text schneller auswählen und machten dabei weniger Fehler. Diese Untersuchung war ausschlaggebend dafür, dass sich die Maus als eines der Haupteingabegeräte des Computers durchgesetzt hat. Sie gilt als eine der ersten klassischen Arbeiten aus der Mensch-Computer-Interaktion.

Seit den 1960er Jahren ist das Grundprinzip der Maus weitestgehend gleich geblieben. Es hat allerdings weitere Verbesserungen der Technologie gegeben, sodass die meisten modernen Mäuse mit optischen Sensoren arbeiten und gegebenenfalls auch kabellos sind.

Optische Mäuse verwenden Fotosensoren, um Strukturen auf der Tischoberfläche zu erkennen und daraus eine Bewegung der Maus abzuleiten. In der Mitte der 1990er Jahre hat außerdem Microsoft angefangen, Scrollräder in Mäusen zu verbauen.

### ✓ Übung: Verschiedene Eingabegeräte

Wir haben bereits viele verschiedene Werkzeuge und Gesetzmäßigkeiten aus der MCI kennen gelernt. Versuchen Sie mithilfe des Gelernten die Eingabegeräte aus diesem Abschnitt zu vergleichen. Können Sie argumentativ bestätigen, dass die Maus das beste Gerät zum Auswählen von Text ist?

**Tipp** Benutzen Sie nicht nur ein Modell, sondern kombinieren Sie zum Beispiel GOMS mit Fitts' Law.

## 6.2.2 Tastatur

Die Ursprünge der Tastatur liegt in der Schreibmaschine. In der Mitte der 1870er Jahre stellte die Firma Remington die erste in Serienfertigung her gestellte Schreibmaschine her. Diese verwendete das QWERTY-Layout, welches sich auch auf Computertastaturen bis heute durchgesetzt hat.

### Tastaturlayout

#### Definition: Tastaturbelegung

Die Tastaturbelegung (aka Tastaturlayout) beschreibt die Kodierung der einzelnen Tasten sowie deren Lage und Häufigkeit auf der Tastatur.

Das Layout einer Computertastatur ist hauptsächlich deswegen so wie es ist, weil es vom Tastaturlayout der Schreibmaschine kopiert wurde. Physikalisch gibt es keinen Grund, dass eine Tastatur so aussehen muss, wie sie es tut. Es ist auf den meisten modernen Betriebssystemen sogar möglich, die Belegung der Tasten einer Tastatur virtuell zu ändern und somit die Tastaturbelegung anzupassen. Bei den Schreibmaschinen ging dies allerdings nicht. Hier musste man sich bei der Konstruktion für ein Layout entscheiden.

**ABC-Layout** Ein naheliegendes Tastaturlayout ist, die Buchstaben nach ihrer Reihenfolge im Alphabet anzuordnen. Um dann noch effizient tippen zu können, kann man anstatt alle Tasten nebeneinander zu setzen die Tasten in mehrere Zeilen aufteilen. Dieses Layout wurde auch für einige Schreibmaschinen angewendet.

Schreibmaschinen schreiben, indem beim Tippen einer Taste ein physikalischer Hebel ausgelöst wird, der auf das Papier schlägt und dort einen Abdruck eines Buchstaben erzeugt. Nach dem Auftreffen des Hebels bewegt sich das Papier ein Stück nach links, damit das nächste Zeichen rechts daneben gesetzt wird. Diese Funktionsweise bewirkt,

dass jede Taste einen eigenen Hebel (mit eigenem Buchstaben) braucht und dass alle Hebel an die Stelle auf dem Papier ausgerichtet sind, an dem der nächste Buchstabe geschrieben werden soll. Dies hatte bei schnellem Tippen zur Folge, dass sich die Hebel verhakten. Insbesondere bei häufig auftretenden Buchstabenkombinationen wie z.B. „DE“ war das Risiko dafür sehr groß.

**QWERTY-Layout** Um das Problem der sich verhakenden Hebel zu minimieren, hat Christopher Sholes Layout entwickelt, bei dem die Buchstaben, die statistisch häufig gemeinsam auftauchen, möglichst weit voneinander entfernt sind (z.B. „QU“). Dieses Layout ist als QWERTY-Layout bekannt geworden (benannt nach den ersten sechs Buchstaben der obersten Reihe auf der Tastatur) und in Abbildung 6.4a zu sehen. Im deutschen gibt es ein sehr ähnliches Layout, das QWERTZ-Layout, das denselben Ansatz verfolgt, aber aus dem statistische Vorkommen der Buchstaben in der deutschen Sprache hervorgeht.

**Dvorak-Layout** Das QWERTY-Layout hat sich vor allem aufgrund von mechanischen Limitationen für Schreibmaschinen durchgesetzt. Bei Computern gibt es diese Einschränkungen nicht mehr. Hier kann man die Tasten beliebig anordnen. Dafür stellen wir als erstes fest, dass das Schreiben mit einer QWERTY-Tastatur nach Fitts' Law (s. Abschnitt 5.5.2) nicht optimal ist, da die Wege zwischen häufig gemeinsam auftretenden Buchstaben sehr groß sind. Man kann aber Fitts' Law auch anwenden, um genau ein Layout zu entwickeln, das für höchste Effizienz ausgelegt ist. Dieses Layout nennt man Dvorak-Layout (nach dem Erfinder August Dvorak). Es ist in Abbildung 6.4b abgebildet.

**Colemac-Layout** Ein Problem beim Dvorak-Layout ist, dass sich die Anordnung der Tasten fundamental von der Anordnung beim QWERTY-Layout unterscheidet. Die meisten Menschen sind aber mit dem QWERTY-Layout vertraut, sodass eine Umstellung auf Dvorak sehr mühsam wäre. Das Colemac-Layout versucht dieses Problem zu lösen, indem es die Vertrautheit von QWERTY mit der Effizienz von Dvorak kombiniert. Dies geschieht, indem nur einige wenige Tasten an andere Plätze bewegt werden, um die größtmögliche Effizienzsteigerung bei der geringsten Umstellung zu erreichen (s. Abbildung 6.4c).

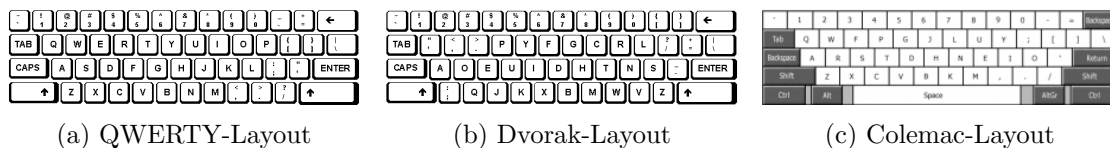


Abbildung 6.4: Tastaturlayouts

### Ergonomische Tastaturen

Seit den 70er Jahren war die Computertechnologie so weit entwickelt, dass es für Privatpersonen möglich wurde, einen eigenen Computer zu besitzen. Computer wurden auch

in Unternehmen zunehmend wichtiger, da sich viele Aufgaben dadurch beschleunigen ließen. Je mehr Menschen an Computern arbeiteten, desto mehr Beschwerden gab es über schmerzende Handgelenke vom Tippen an Tastaturen. Dies brachte in den 80er und 90er Jahren den Bereich Human Factors und Softwareergonomie hervor. Diese Forschungsgebiete beschäftigen sich mit dem Design von Computerschnittstellen (in der Hard- und Software) auf Grundlage der physiologischen Eigenschaften des Menschen. In der Folge wurden ergonomische Tastaturen entwickelt, die die natürliche Handhaltung des Menschen berücksichtigen und ein angenehmeres Layout bieten sollen. Hier tritt allerdings dasselbe Problem wie beim Dvorak- und Colemac-Layout auf: Den meisten Menschen ist die Umgewöhnung an ein neues Tastaturlayout zu aufwändig.

Blatt 14 - Ergonomie

Ergonomische  
Tastatur

### **Virtuelle Tastaturen**

Die Entwicklung von Smartphones und Tablet-Computern hat einige wichtige Neuerungen im Bereich Tastaturen mit sich gebracht. Auch wenn solche virtuellen Tastaturen oft weiterhin auf dem QWERTY-Layout der Schreibmaschinen basieren, haben sie deutliche Vorteile gegenüber physischen Tastaturen. Durch ihren virtuellen Charakter ist es möglich, das Tastaturlayout dynamisch anzupassen. Das geht sogar so weit, dass abhängig vom Kontext verschiedene Tastaturlayouts verwendet werden, die es dem Benutzer erleichtern, bestimmte Daten einzugeben. So werden bei einem Feld, das nur Zahlen als Eingabe erlaubt, auch nur Zahlen auf der Tastatur angezeigt.

# 7 Interaktionsparadigmen

## Definition: Interaktionsparadigma

Ein Interaktionsparadigma beschreibt die Art und Weise, wie mit einem Computer interagiert wird. Die Interaktion erfolgt über eine Benutzerschnittstelle (s. Abschnitt 2.1).

Im Abschnitt 2.1 haben wir bereits die Benutzerschnittstelle als Mittler zwischen Mensch und Computer kennen gelernt. In diesem Kapitel werden wir detaillierter auf verschiedene Klassifikationen von Benutzerschnittstellen und die dazugehörigen Interaktionsparadigmen eingehen und wichtige Begriffe kennen lernen, mit denen sich Benutzerschnittstellen beschreiben lassen. Der größte Teil dieses Kapitels wird sich auf Monitor-basierte Benutzerschnittstellen beziehen. Das hat den Grund, dass diese Art der Benutzerschnittstelle die häufigste bei der Interaktion mit Computern ist. Am Ende des Kapitels werden wir aber auch einige Schnittstellen sehen, die sich fundamental von den klassischen Interaktionen mit Computern unterscheiden.

## 7.1 Klassische Interaktionsgestaltung

### 7.1.1 Kommandosysteme

Kommandosysteme (*engl. Command Line Interface, CLI*) bieten eine textbasierte Interaktion durch eine Menge von vordefinierten Kommandos. Normalerweise können die Kommandos durch optionale oder Pflichtparameter gesteuert werden. Ein Beispiel für ein solches System ist BSD-Unix mit 438 verschiedenen Kommandos. Kommandozeilensysteme werden heutzutage vor allem von Experten verwendet und besonders in solchen Situationen eingesetzt, in denen variable, komplex zusammenhängende Funktionen benötigt werden.

**Das Paradigma** Kommandosystem basiert auf deskriptiven Interaktionsformen (sprachlichen Beschreibungen) und erfordert das **Erinnern** an Befehlsnamen, deren Syntax sowie Optionen und Parameter. Kommandosysteme sind im Vergleich zu anderen Paradigmen sehr effizient und bieten häufig mehr Funktionalitäten. Dies gilt besonders, da der Funktionsumfang durch das Definieren neuer Kommandos praktisch unbegrenzt erweitert werden kann.

### 7.1.2 Form-Fill Interfaces

Form-Fill Interfaces bestehen aus drei Elementen:

1. Ein Formular
2. Felder auf dem Formular für die Eingabe von Daten
3. Möglichkeiten zur Navigation durch die Felder sowie weitere Eingabemöglichkeiten

Form-Fill Interfaces werden sehr häufig eingesetzt, wenn es darum geht, Daten einzugeben. Ein typisches Beispiel für solche Interfaces sind Tabellenkalkulationsprogramme oder Formulare auf Websites. Beide zeichnen sich dadurch aus, dass es Felder gibt, in denen Daten eingetragen werden können. Der Vorteil dieses Paradigmas gegenüber Kommandosystemen ist vor allem, dass die Dateneingabe auch für Nicht-Experten deutlich vereinfacht wird. Allerdings geschieht dies auf Kosten des Funktionsumfangs und des Bildschirmplatzes.

**Das Paradigma** Form-Fill Interface minimiert durch das Anbieten von Auswahllisten oder dynamisch generierten Vorschlägen den Erinnerungsaufwand und erlaubt das **Wiedererkennen**. Auch Form-Fill Interfaces basieren eher auf deskriptiven Interaktionsformen.

### 7.1.3 Point & Click Interfaces

Point & Click Interfaces zeichnen sich durch Kommando- oder Optionselemente aus

1. auf welche mit einem Zeigegerät (z.B. Maus oder Finger) gezeigt wird
2. die selektiert sowie angeklickt werden können

Das Ziel dieses Interaktionsparadigma ist es, die nötigen Tastatureingaben zu minimieren. Dazu werden sie meistens mit Menu Interfaces (s.u.) kombiniert, um die nötige Menge an Optionen anzubieten. Die meisten modernen Betriebssysteme mit grafischer Benutzeroberfläche basieren auf Point & Click Interfaces mit der Maus als Zeigegerät. Gleiches gilt auch für mobile Betriebssysteme wie iOS oder Android, bei denen über Touch-Eingaben mit dem Finger gezeigt und geklickt wird.

Besonders die Interaktion über einen Touchscreen ist für den Menschen sehr intuitiv und wird daher auch häufig verwendet (z.B. bei Fahrkartenautomaten oder auf Websites). Da die Anzahl der möglichen Operationen durch den Bildschirmplatz begrenzt ist, treten bei der Interaktion mit einem solchen Interface auch oft weniger Fehler auf. Allerdings gilt auch hier, dass der Funktionsumfang gegenüber Kommandozeileninterfaces eingeschränkt ist. Außerdem ist das physische Zeigen im Vergleich zum Tippen auf der Tastatur deutlich aufwändiger und auch weniger effizient, sodass das Erfüllen einer Aufgabe mit einem Point & Click Interface mehr Zeit in Anspruch nimmt, als das Erfüllen einer ähnlichen Aufgabe mit einer Kommandozeile.

**Das Paradigma** Point & Click Interface verwendet eher deiktische Interaktionsformen (Selektion mittels Zeigerhandlung) und erlaubt somit das **Wiedererkennen** von Elementen. Das Design eines guten Point & Click Interfaces, bei dem möglichst wenig Mausebewegungen erforderlich sind, ist sehr aufwändig. Im Kapitel Interaktionsdesign werden wir einige Techniken kennen lernen, die beim Design eines guten Interfaces helfen. Der Vorteil an einem gut gestalteten Interface ist allerdings, dass damit auch semi-erfahrene gut umgehen können.

#### 7.1.4 Menu Interfaces

Menu Interfaces basieren auf Menüstrukturen. Ein Menü ist eine Menge von Optionen auf dem Bildschirm, wobei die Selektion und Ausführung einer oder mehrerer Optionen zur Veränderung eines Zustands führen. Da die Auswahl der Optionen im Menü in der Regel über Point & Click Mechanismen erfolgt, erben Menu Interfaces normalerweise die Vor- und Nachteile von Point & Click Interfaces. Es ist allerdings auch möglich, eine Menüauswahl mithilfe der Tastatur vorzunehmen.

**Das Paradigma** Menu Interface verwendet eher deiktische Interaktionsparadigmen, da die einzelnen Optionen klassischerweise mit der Maus angeklickt werden. Dadurch, dass aus einer vorgegebenen Menge an Optionen ausgewählt wird, erlauben diese Interfaces außerdem das **Wiedererkennen**. Wie auch bei Point & Click Interfaces ist eine gute Strukturierung von Menu Interfaces sehr wichtig. Zu tiefe Verschachtelungen erhöhen einerseits den Kognitiven Aufwand, da sich der Benutzer an den Ort des Eintrags erinnern muss, und erhöhen andererseits nach Fitts' Law die Zeit, die der Benutzer zum Anklicken einer Option braucht.

#### 7.1.5 Direkte Manipulation

Diese Interfaces sind ein Spezialfall von Point & Click Interfaces, bei dem der Benutzer ein Objekt durch Mausoperationen direkt beeinflussen kann und ein kontinuierliches Feedback über die Auswirkungen seiner Aktionen bekommt. Meistens wird diese Art Interface für Design-Programme verwendet (z.B. zur 3D-Modellierung).

Ein häufig angewendetes Hilfsmittel bei Direkter Manipulation sind Interaktionsmetaphern. Bei Interaktionsmetaphern wird eine aus dem Alltag bekannte Beziehung von Aktion und Reaktion auf eine interaktive Anwendung übertragen. Führt man dies weiter, erhält man ganze Metaphorische Dialoge. Bei Direkter Manipulation arbeitet man sehr viel auf dieser metaphorischen Ebene, da die meiste Werkzeuge mit Interaktionsmetaphern benannt sind. So gibt es bei Bildverarbeitungsprogrammen oft einen Pinsel, Radiergummi oder Farbeimer. Besonders bei generischen Operationen wie „auswählen“ oder „vergrößern“ findet man auch oft Metapher, die in verschiedenen Programmen gleich verwendet werden. Andere bekannte Interaktionsmetaphern sind zum Beispiel „Dokument“, „Ordner“ oder „Einkaufswagen“.

Direkte Manipulation hat gegenüber anderen Interaktionsparadigmen vor allem den Vorteil, dass durch Interaktionsmetaphern wenig neue Begriffe und Operationen erlernt

Interaktionsmetaphern

Metaphorische  
Dialoge

werden müssen, bzw. dass die Operationen sehr schnell verständlich sind. Durch das sofortige Feedback und die Möglichkeit Aktionen zu widerrufen liegt die Kontrolle bei diesem Interaktionsparadigma klar beim Benutzer selbst. Allerdings gilt auch hier im Vergleich zu Kommandosystemen, dass bei der Komposition komplexer Funktionen viel Zeit durch Point & Click Operationen verschwendet wird. Im Bereich der Softwareentwicklung ist das Entwickeln von Programmen, die Direkte Manipulation erlauben, ebenfalls deutlich aufwändiger.

**Das Paradigma** Direkte Manipulation zeichnet sich durch eine permanente Darstellung des zu manipulierenden Objektes und sofortiges Feedback aus. Dabei werden hauptsächlich deiktische Interaktionsformen angewendet, wodurch auch hier **Wiedererkennen** anstelle von Erinnern möglich ist.

## 7.2 WIMP-GUIs

### Definition: Grafische Benutzerschnittstelle

Eine Grafische Benutzerschnittstelle (*engl. Graphical User Interface, GUI*) erlaubt die Interaktion des Menschen mit dem Computer über grafische Symbole.

Im Allgemeinen können GUIs sehr unterschiedlich aussehen. Jede Form von Benutzerschnittstelle, die mehr als nur Text anzeigt, kann in gewisser Weise als grafische Benutzerschnittstelle angesehen werden. Bestimmte Elemente findet man jedoch bei sehr vielen grafischen Benutzerschnittstellen wieder. Dies führt zur Definition der WIMP-GUIs:

### Definition: WIMP-GUI

WIMP (**W**indows, **I**cons, **M**enus, **P**ointer) sind spezielle Symbole bzw. Window Gadgets (abk. Widgets), die bei der Gestaltung einer GUI eingesetzt werden.

Der erste Computer, der WIMP-GUIs verwendete war der Xerox Alto, der 1977 am Xerox Parc vorgestellt wurde. Dieser inspirierte Steve Jobs zur Entwicklung des Apple Lisa, was wiederum den Grundstein für die Entwicklung des Macintosh sowie des Windows-Betriebssystems legte. Der erste kommerziell erhältliche Computer mit GUI war der Xerox SDD mit dem Star Interface, der ab 1981 erhältlich war. Schon in diesen ersten GUI-Rechnern waren viele WIMP-Elemente wiedererkennbar.

### 7.2.1 Windows

Fenster sind das erste Element von WIMP-GUIs. Fenster werden verwendet, um das Problem des eingeschränkten Bildschirmplatzes zu lösen. Bevor es Fenster gab, nahm ein geöffnetes Programm den kompletten Bildschirm ein. Die Einführung von Fenstern (insbesondere auch in Kombination mit Scrollbars) erweiterte den physikalischen Bildschirm



auf praktisch unbegrenzt viel Platz. Dadurch, dass außerdem mehrere Fenster gleichzeitig angezeigt werden können, wird die Menge an Informationen, die gleichzeitig auf dem Bildschirm angezeigt werden können deutlich erhöht.

Die Erfindung von Fenstern zur Erweiterung des physikalischen Bildschirms geht bis zur „Mother of all Demos“ zurück. Damals waren die Fenster allerdings noch fest auf dem Bildschirm verankert und man konnte sie nicht verschieben. Erst am in den Jahren 1969-1974 wurde die Funktionalität von überlappenden Fenstern erforscht und letztlich auch in kommerziellen Systemen eingesetzt. 1984 wurde dann das X Windows System am MIT entwickelt, das bis heute die Grundlage für Fenster auf Unix-basierten Systemen bildet.

**Pop-Ups** Pop-Ups sind spezielle Arten von Fenstern, die zusätzliche Inhalte anzeigen und/oder die Interaktion des Benutzers benötigen. Oft muss mit diesen Fenstern interagiert werden, bevor mit anderen Fenstern des Programms interagiert werden kann. Man spricht hier von modalen Fenstern. Pop-Ups nehmen den Benutzer aus einem Arbeitskontext heraus und führen sehr leicht zur Ablenkung. Daher sollte man Pop-Ups nur selten und sehr bewusst einsetzen.

Modales Fenster

## 7.2.2 Icons

Icons sind ikonische Repräsentationen von Objekten, die in einer WIMP-GUI angezeigt werden. Icons werden häufig zusammen mit der Desktop-Metapher verwendet und schaffen so einen virtuellen Raum, mit dem der Benutzer interagieren kann. Auch andere Interaktionsmetaphern werden in grafischer oder Textueller Form bei Icons eingesetzt. Beispiele dafür sind „Dokumente“ oder „Ordner“.

Bei der Gestaltung von Icons spielen viele Faktoren eine Rolle. Neben der Form des Icons und einer eventuell abgebildeten Grafik können verschiedene Genres von Objekten ebenfalls durch visuelle Eigenschaften (Kontrast, Farbe, Salienz etc.) getrennt werden. So sind bei macOS Icons von Hilfsprogrammen tendentiell eher gräulich, während Icons von normalen Anwendungen sehr farbenfroh sind.

## 7.2.3 Menus

Menüs haben wir bereits bei den Menu Interfaces kennen gelernt. Ein Menü ist eine Struktur zur Auswahl aus einer geordneten Menge von Optionen. Die Optionen sind häufig hierarchisch angeordnet. Bei WIMP-GUIs gibt es verschiedene Arten von Menüs:

- Drop-down Menüs
- Kontextmenüs
- Verschachtelte Menüs
- Scrollmenüs
- Tortenmenüs

- ...

Jedes dieser Menüs hat Vor- und Nachteile in bestimmten Situationen. In Kapitel 8 beschäftigen wir uns eingehender mit Techniken zum auswählen und Anordnen von Menüs.

#### **7.2.4 Pointer**

Das letzte Element einer WIMP-GUI ist der Pointer. Der Pointer ist ein grafisches Bild auf dem Display, welches das Zeigegerät (z.B. die Maus) und dessen Bewegungen repräsentiert. Der Pointer kann dann benutzt werden, um Icons auszuwählen oder mit Widgets zu interagieren.

#### **7.2.5 Widgets**

Widgets (Window Gadgets) sind die zentralen Komponenten, aus denen sich eine WIMP-GUI zusammensetzt. Widgets definieren einen Bereich in einem Fenster, der auf Maus- und/oder Tastaturereignisse reagiert und als Reaktion darauf Zustände der Komponente verändern kann. Die klassischen Widgets sind z.B. Buttons, Checkboxes, Textfelder, Slider, ...

Insbesondere Buttons sind für WIMP-GUIs sehr relevant. Ein Button ist ein individueller, isolierter und selektierbarer Display-Bereich, welcher Aktionen oder Zustandsänderungen hervorruft. Jedes Widget, das auf eine Selektion reagiert, kann als Button klassifiziert werden. Dazu zählen z.B. Checkboxes, Radio-buttons, Spin-Buttons, ...

#### **7.2.6 Look-&-Feel**

Das WIMP-Paradigma definiert die grundsätzlichen Elemente einer GUI sowie das grundsätzliche Verhalten der einzelnen Widgets. Es definiert nicht das genaue Aussehen der Elemente. Alle nicht-funktionalen Eigenschaften der Widgets einer WIMP-GUI nennt man „Look-&-Feel“. Das Look-&-Feel umfasst z.B. die Größe, Farbe, Schritart und viele andere Eigenschaften der Widgets aber auch das spezielle Verhalten. Zum Beispiel könnte ein Look-&-Feel festlegen, dass eine Option in einem Menü aktiviert wird, wenn die Maustaste gedrückt wird, während ein anderes Look-&-Feel die Option erst aktiviert, wenn die Maustaste losgelassen wird.

### **7.3 Weitere Interface-Type**

Neben WIMP-GUIs gibt es noch viele weitere Arten von Benutzerschnittstellen. Diese kommen allerdings nur selten auf traditionellen Desktop-Computern zum Einsatz, sondern werden meistens auf anderen Computersystemen verwendet. Wir unterscheiden die folgenden Typen von Interfaces:

1. Kommandosysteme
2. WIMP & GUI

3. Multimedia
4. Virtuelle Realität
5. Informationsvisualisierung
6. Web-basiert
7. Konsumer-Elektronik & -Geräte
8. Mobile-basiert
9. Sprach-basiert
10. Pen-basiert
11. Touch-basiert
12. Gesture-basiert
13. Haptik
14. Multimodalität
15. Shareable
16. Taggible
17. Erweiterte & Gemischte Realität
18. Wearable
19. Robotik
20. Brain-Computer

Wir werden im Folgenden auf einige der Interface-Typen exemplarisch eingehen. Viele der Aspekte, die bereits bei WIMP eine Rolle gespielt haben, gelten auch für andere Interface-Typen (z.B. Fitts' Law und Hick's Law). Insbesondere bleiben die Rahmenbedingungen, die durch das Model Human Processor gegeben sind, weiter gültig.

**Konsumgeräte** Konsumgeräte (*engl. Consumer-Devices*) sind Geräte des alltäglichen Gebrauchs. Dazu zählen Waschmaschinen, Fernbedienungen, Navigationssysteme usw. Im Gegensatz zu Desktop-Computern, die viele verschiedene Aufgaben erfüllen müssen, liegt der Fokus bei Interfaces für Konsumgeräte darauf, dass eine bestimmte Aufgabe möglichst schnell erledigt werden kann. Man muss dabei beachten, dass viele Konsumgeräte nur sehr geringe Kapazitäten haben. Sowohl die Rechenleistung als auch der Speicher sind sehr begrenzt. Zusätzlich ist der Monitor noch sehr klein.

Das Ziel bei der Entwicklung von Interfaces für Konsumgeräte ist, jedem Nutzer das Interface sofort verständlich zu machen. Um dies zu erreichen wird im Gegensatz zu WIMP häufig auf Software-Buttons verzichtet und es werden stattdessen physische Knöpfe und Rädchen verbaut.

**Mobile User Interfaces** Das Besondere an dieser Art von Interface ist, dass man den Computer prinzipiell immer bei sich tragen kann. Das erlaubt neue Arten von Software, die z.B. ortsabhängige Interaktionen ermöglichen.

**Sprach-basierte UIs** Sprach-basierte UIs verwenden den natürlichsten und gebräuchlichsten Kanal der menschlichen Kommunikation, um eine *echte* Kommunikation zwischen Mensch und Computer zu ermöglichen. Eine vollständige sprach-basierte UI zeichnet sich durch symmetrische Modalität aus:

**Synthese** Der Computer kann Texte digital abbilden und in Audio verwandeln

**Analyse** Der Computer versteht die Sprache des Benutzers

Heutige Sprach-basierte UIs (wie zum Beispiel Siri) sind schon ziemlich gut beim Erkennen natürlicher Sprache. Die Erkennungsquoten liegen bei ca. 95 % oder sogar noch höher. Da Sprache aber ein so intuitives Kommunikationsmittel für den Menschen ist, fallen Fehler des Computers sehr schnell auf. Ein anderes Problem ist, dass die Korrektur eines Fehlers in den meisten Fällen die Interaktion mit einer nicht sprach-basierten UI erfordert (z.B. Tastatureingabe). Daher haben sich sprach-basierte UIs noch nicht als Standard für die Interaktion zwischen Mensch und Computer durchgesetzt.

**Pen-basierte UIs** UIs mit Stifteingabe sind eng verwandt mit Touch-basierten Interfaces. Der Vorteil eines Eingabestiftes ist, dass die Interaktion deutlich präziser ist. Diese Art von UIs werden hauptsächlich in Situationen benutzt, in denen auch in der Realität ein Stift benutzt werden würde (z.B. zum Zeichnen oder Schreiben).

**Touch-basierte UIs** Touch-basierte UIs verwenden als Zeigeobjekt den Finger des Menschen. Sie werden häufig in Kombination mit anderen Interface-Typen verwendet (z.B. Mobile User Interfaces oder Konsumgeräte). Es gibt einige unterschiedliche Touchscreen-Technologien (optisch, resistiv, kapazitiv, ...), allerdings hat die Art des Touchscreens keine direkte Auswirkung auf die UI.

**Air-basierte UIs** Air- oder Mid-Air-basierte UIs sind Benutzerschnittstellen, die Gesten im Raum erkennen und interpretieren. Sie ähneln Touch-basierten UIs in der Hinsicht, dass bei beiden Schnittstellen die Hände als Zeigewerkzeug benutzt werden. Bei Air-basierten UIs fehlt allerdings jegliches haptisches Feedback. Ein Beispiel für so eine UI ist das Microsoft Kinect-System, die Gesten durch *Structured Light* erkennt.

**Haptische UIs** Genau gegenteilig zu Air-basierten UIs arbeiten haptische Benutzerschnittstellen. Hier steht die Haptik im Mittelpunkt. Im Gegensatz zu Touchscreens wird aber der Berührungssinn nicht nur durch Berührung sondern auch durch Kraft, Vibration oder Bewegungen ausgenutzt. Die Ein- und Ausgabe erfolgt dabei in der Regel über dasselbe Gerät. Bei der Eingabe bewegt der Benutzer ein Eingabegerät, bei der Ausgabe bewegt sich das Gerät von selbst (oder blockiert die Bewegung).

**Multimodale UIs** Multimodale UIs kombinieren mehrere Modalitäten des Benutzers (z.B. Sprache und Gesten) und unterstützen Eingaben, die aus multisensorischen Daten kommen. In der Regel liegt auch bei solchen Systemen eine symmetrische Modalität vor, das heißt Benutzer und Computer können über dieselbe Modalität kommunizieren.

**Tangible UIs** Diese Schnittstellen ermöglichen die Interaktion mit digitalen Informationen über eine physikalische Umgebung. Diese Art von UI ist besonders für das Paradigma Direkte Manipulation (s. Abschnitt 7.1.5) interessant. Die Idee ist hier, dass man ein physikalisches Objekt verformt und automatisch eine entsprechende digitale Repräsentation erhält. Dies gilt genauso in umgekehrter Weise, dass also digitale Transformationen direkt in der physikalischen Welt zu sehen sind.

**Mixed Reality** Unter den Begriff Mixed Reality fallen zwei Unterkategorien, die eine Verbindung zwischen der echten Welt (*engl. Reality*) und einer virtuellen Welt (*engl. Virtual Reality*) herstellen.

**Augmented Reality** Überlappung der realen Welt durch virtuelle Informationen.

**Augmented Virtuality** Erweiterung der virtuellen Welt durch Objekte oder Informationen aus der realen Welt.

Diese Begriffe sind auch in Abbildung 7.1 dargestellt.

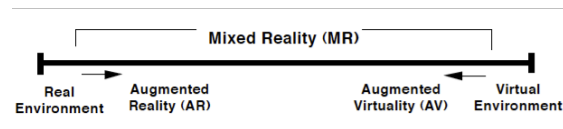


Abbildung 7.1: Mixed Reality

**Immersive Interfaces** Immersive Interfaces sind Benutzerschnittstellen, die den Benutzer in eine virtuelle Welt projizieren und dem Menschen auch das Gefühl geben, tatsächlich in der virtuellen Welt präsent zu sein.

**Roboter-basierte UIs** Diese Interfaces ermöglichen die Interaktion des Menschen mit dem Computer über Roboter, die normalerweise menschliche Gestalt haben. Zu dieser Art der UI zählt auch die Art der Fernsteuerung von Robotern, bei der einem suggeriert wird, dass man sich im Roboter befindet. Das Forschungsfeld der Human-Robot-Interaction beschäftigt sich mit der Entwicklung von solchen Robotern und UIs.

# 8 Interaktionsdesign

## Definition: Interaktionsdesign

Interaktionsdesign befasst sich mit der Gestaltung der Funktion, des Verhaltens und der endgültigen Ausgestaltung von Produkten und Systemen.

Interaktionsdesign findet auf drei essentiellen Ebenen statt und begleitet den gesamten Entwicklungsprozess eines Produktes. Die verschiedenen Ebenen des Interaktionsdesign laufen teilweise sequentiell oder parallel und können sich gegenseitig beeinflussen. Bei der Entwicklung eines Produktes müssen am Ende alle Ebenen zueinander passen, um einen passenden Gesamteindruck zu erzeugen. Die Ebenen sind:

- Hardwaredesign
- Softwaredesign
- Mediendesign

## 8.1 Leitsatz für Gutes Design

*„Whether it be the sweeping eagle in his flight  
or the open apple-blossom  
the toiling work-hose,  
the blithe swan,  
the branching oak,  
the winding stream at its base,  
the drifting clouds  
over all the coursing sun,  
form ever follows function, and this is the law.“*

— L. Sullivan: The Tall Office Building Artistically Considered in Lippincott's Magazine, 1896

Der Kernpunkt des Leitsatz ist die Aussage „Form folgt Funktion“, das heißt die Form soll sich aus der Funktion ableiten und umgekehrt soll sich die Funktion aus der Form ableiten lassen. Dahinter steht der Gedanke, dass das Design einem bestimmten Zweck dient und die Form dabei helfen soll, diesen Zweck zu erfüllen. Wird der Leitsatz beim Design eines Produktes eingehalten, kann das Produkt einfach und zielführend benutzt werden.

**Ästhetik** Beim Leitsatz von Sullivan ist sehr wichtig, dass auch Ästhetik als eine Funktion gilt, der die Form folgt. Dies führt oft zu Missverständnissen des Grundsatzes und zur Auslegung, dass die *Schönheit* einer Sache irrelevant ist. Tatsächlich hat aber auch Sullivan selbst bei dem Haus, das er gebaut hat, Ornamente und Verzierungen angebracht.

Bei interaktiven Systemen oder Produkten stehen Nutzer, Eigentümer oder Betroffene beim Design im Vordergrund. Das Design eines solchen Systems ist *gelingen*, wenn das System einen Zweck erfüllt, das heißt die Menschen erreichen mit dem System oder Produkt ihre Aufgaben und Ziele. Um dies zu unterstützen, gibt es einige Gestaltungskriterien, die den Designprozess leiten und die Ergebnisse bewertbarer machen. Diese Kriterien sind

1. Funktionalität der Gestaltung
2. Ergonomie der Gestaltung
3. Ästhetik der Gestaltung
4. Erlebnishaftigkeit der Gestaltung
5. Symbolik der Gestaltung

## 8.2 Affordanzen

### 8.2.1 Reale Affordanzen

Neben den obigen Designkriterien hat ein Design insbesondere eine funktionale Bedeutung. Das heißt, dass das Erscheinungsbild eines Objektes auch vermittelt, wie man dieses Objekt benutzt. Affordanzen beschreiben, wie über eine Erscheinungsbild die Funktionsweise eines Objektes vermittelt wird:

#### Definition: Affordanzen (nach Gibson)

*„Affordances are all action possibilities latent in the environment, objectively measurable and independent of the individual's ability to recognize them, . . .“*

— J.J. Gibson: The Theory of Affordances. In *Perceiving, Acting, and Knowing*, Eds. Robert Shaw and John Bransford, 1977

Affordanzen nach Gibson sind also alle Gebrauchseigenschaften eines Objektes, die objektiv messbar sind. Der Benutzer spielt hier explizit keine Rolle. Im Gegensatz dazu steht eine alternative Auslegung von Norman:

### Definition: Affordanzen (nach Norman)

*„... the term affordance refers to the perceived and actual properties of the thing, primarily those fundamental properties that determine just how the thing could possibly be used... Affordance provides strong cues to the operation of things.“*

— Don Norman (2013). The Design of Everyday Things, p. 11

Nach Norman fallen also nur solche Gebrauchseigenschaften unter den Begriff „Affordanz“, die auch vom Benutzer wahrgenommen werden. Tabelle 8.1 stellt noch einmal die Unterschiede der beiden Definitionen dar.

#### Affordanzen nach Gibson

- Action possibilities in the environment in relation to the action capabilities of an actor.
- Independent of the actor's experience, knowledge, culture, or ability to perceive.
- Existence is binary - an affordance exists or it does not exist.

#### Affordanzen nach Norman

- Perceived properties that may not actually exist.
- Suggestions or clues as to how to use the properties.
- Can be dependent on the experience, knowledge, or culture of the actor.
- Can make an action difficult or easy.

Tabelle 8.1: Affordanzen nach Gibson und Norman

In der Mensch-Computer-Interaktion und beim Interaktionsdesign interessieren uns vor allem diejenigen Gebrauchseigenschaften, die vom Benutzer wahrgenommen werden und auch durchgeführt werden können (Erkennbare Affordanzen). Diese Affordanzen sind genau die Schnittmenge zwischen den Affordanzen von Gibson und Norman (Gibsons Affordanzen sind die existierenden und Normans die wahrgenommenen). Bei Gebrauchseigenschaften, die nicht in den Perceptible Affordances liegen, spricht man von Falsche Affordanzen (im Falle von wahrgenommenen, nicht existierenden Affordanzen), bzw. Versteckte Affordanzen (Im Falle von existierenden, nicht wahrgenommenen Affordanzen).

Erkennbare  
Affordanzen

Falsche Affordanzen

Versteckte  
Affordanzen

Das Ziel beim Design eines Objektes oder einer UI ist es nun, eine möglichst große Schnittmenge der erkennbaren Affordanzen zu erhalten. Es gibt keine festen Regeln dafür, wie eine Affordanz erkannt wird. Hauptsächlich spielt die Lebenserfahrung des Benutzers eine Rolle. Allerdings gibt es einige Einschränkungen, die man sich im Bezug auf Affordanzen zunutze machen kann:

**Physische Limitierungen** z.B. schwere Geräte können vom Menschen nicht benutzt werden.

**Physikalische Limitierungen** z.B. Mauszeiger kann nicht über den Bildschirmrand hinaus bewegt werden.



**Logische Limitierungen** z.B. Vorgabe für Anordnung von Schalterleisten für Geräte

**Kulturelle Limitierungen** z.B. Rote Ampel bedeutet „Stopp“.

## 8.2.2 Virtuelle Affordanzen

In der virtuellen Welt des Computers gelten die physischen und physikalischen Einschränkungen der echten Welt nicht mehr. Trotzdem können virtuelle Objekte dem Benutzer Affordanzen vermitteln (z.B. durch Form und Farbe). Diese Affordanzen orientieren sich entweder an der Erfahrung des Nutzers oder basieren auf ähnlichen Objekten in der echten Welt. Insbesondere Schattierungen sind ein beliebtes Mittel, um den Realitätsbezug herzustellen (z.B. bei Buttons).

### Aufgabe 23: Affordanzen

Ordnen Sie die folgenden Textausschnitte den Kategorien von Affordanzen zu (die Affordanz soll dabei das Verhalten des Textes als Hyperlink sein). Nehmen Sie an, dass Sie diese Texte so auf einer Website oder in einem PDF-Dokument finden:

- [Dies ist ein Link](#)
- Dies ist ein Link
- [Dies ist kein Link](#)

**Als Übung** Welche weiteren Affordanzen haben die verschiedenen Texte?

## 8.3 Mentale & Konzeptionelle Modelle

### 8.3.1 Mentale Modelle

In diesem Abschnitt beschäftigen wir uns mit der Frage, wie man Affordanzen einsetzen kann, um eine besonders benutzerfreundliche UI zu designen.

#### **Definition: Mentales Modell**

Ein mentales Modell ist eine Repräsentation eines Gegenstandes oder Prozesses im Bewusstsein eines Lebewesens und ist daher ein Ausschnitt der Wirklichkeit.

Im mentalen Modell eines Gegenstandes ist nicht nur dessen Form und Farbe enthalten, sondern alle Eigenschaften, die man von dem Gegenstand wahrgenommen hat. Auch die Affordanzen des Gegenstandes gehören dazu. Ein Mentales Modell kann einem Benutzer dabei helfen, seine Vorstellung von Gegenständen oder Arbeitsabläufen zu manifestieren. Einige Beispiele für mentale Modelle sind:

- Eine Schraube kann durch das Drehen gegen den Uhrzeigersinn gelöst werden.

- Beim Wasserhahn bedeutet rot warm und blau kalt.
- Das Einstecken eines HDMI-Kabels aktiviert ein externes Display.

Für alle drei angeführten Beispiele kann man auch Gegenbeispiele finden. Das ist aber kein Widerspruch zum mentalen Modell. Ein mentales Modell umfasst nicht alle Fälle, sondern ist viel mehr die Art und Weise, wie sich der Benutzer die Interaktion mit einem Objekt vorstellt (auch wenn das ggf. nicht stimmt).

Mentale Modelle sind die Grundlage jeder Interaktion mit einem unbekanntem Objekt. Sie existieren auch für interaktive Systeme und Produkte. Begegnet man einem unbekanntem Objekt, wird dazu ein mentales Modell auf Grundlage von vorhergegangenen Wissen und Erfahrungen sowie aus der Kombination bereits existierender mentaler Modelle erstellt. Sobald ein Modell existiert, leitet dies dann die Interaktion mit dem Objekt. Sieht man beispielsweise das erste Mal eine virtuelle Tastatur auf einem Touchscreen, leitet das mentale Modell einer physischen Tastatur die Interaktion mit der virtuellen Tastatur. Allerdings funktionieren virtuelle Tastaturen zum Teil anders als physische Tastaturen, sodass das Modell hier an eine Grenze stößt. Um dieses Problem zu lösen, wird das mentale Modell von Tastaturen entsprechend aktualisiert. Dieser Kreislauf ist auch in Abbildung 8.1 dargestellt.

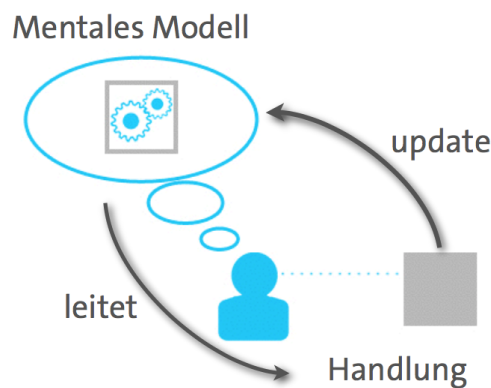


Abbildung 8.1: Mentale Modelle

### 8.3.2 Konzeptionelle Modelle

Die mentalen Modelle verschiedener Menschen können sich stark unterscheiden. Im Interaktionsdesign geht es daher darum, möglichst klare mentale Modelle zu schaffen, die die Arbeit mit einem interaktiven System erleichtern. Ein Werkzeug dafür ist das konzeptionelle Modell.

#### Definition: Konzeptionelles Modell

Ein konzeptionelles Modell beschreibt Konzepte aus Sicht des Anwendungsbereiches

und dient zur Klärung der Terminologie und Anforderungen. Damit spiegelt ein konzeptionelles Modell ein System aus Sicht des Interaktionsdesigners wieder.

Die Entwicklung eines konzeptionellen Modells hift sehr stark dabei, ein konsistentes Erscheinungsbild zu schaffen, was wiederum dazu führt, dass Benutzer ein sehr klares mentales Modell von dem System bekommen. Ein konzeptionelles Modell zu entwickeln ist aber auch deutlich aufwändiger als ein mentales Modell zu entwickeln. In der Regel muss man dazu viele Benutzerbefragungen durchführen und sich mit anderen Menschen austauschen, um so möglichst viele mentale Modelle zu erhalten, aus denen man ein gemeinsames konzeptionelles Modell extrahieren kann. Im Idealfall führt dieser Prozess dazu, dass die mentalen Modelle des Benutzers sofort mit dem konzeptionellen Modell des Entwicklers übereinstimmt. In der Realität ist meistens ein Lernprozess nötig, bei dem die beiden Modelle aufeinander abgebildet bzw. übersetzt werden müssen.

## Interface Metaphern

Um ein möglichst gutes konzeptionelles Modell zu entwickeln (also eines, bei dem das mentale Modell von Anfang an sehr gut übereinstimmt) empfiehlt es sich, Interface Metaphern einzusetzen.

### Definition: Interface Metapher

Interface Metaphern bestehen aus einer Menge von Objekten, Aktionen oder Prozeduren, die das konzeptionelle Modell widerspiegeln.

Oft verwendet man als Metaphern Analogien zur echten Welt (z.B. Schreibtisch oder Papierkorb), die dem Benutzer helfen, ein entsprechendes mentales Modell aufzubauen. Besonders gut sieht man dies bei E-Commerce Systemen. Diese orientieren sich beim konzeptionellen Modell an der Idee eines Selbstbedienungsmarktes. Entsprechend gibt es Interface Metaphern wie „Einkaufswagen“ und „Zur Kasse gehen“. Als Benutzer erstellt man sich dadurch sofort ein mentales Modell auf Grundlage von echten Einkaufswagen und Kassen und versteht dadurch, wie das System zu bedienen ist.

### 8.3.3 Mappings

Die Hauptaufgabe des Interaktionsdesigns ist es, (interne) Intentionen des Benutzers auf (externe) Gestaltungselemente (z.B. Formen, Buttons, Icons) abzubilden. Diese Abbildung nennt man Mapping. Gutes Design zeichnet sich dadurch aus, dass das Mapping von den Intentionen des Benutzers auf die Benutzeroberfläche gelungen ist. Ein ideales Mapping erhält man, wenn man durch gutes Interaktionsdesign dafür sorgt, dass das mentale Modell des Nutzers gut auf das konzeptionelle Modell des Entwicklers abgebildet wird.

Mappings sorgen dafür, dass das Wissen zur Problemlösung, das der Nutzer hat, auf Interaktionselemente abgebildet wird. So braucht der Benutzer nur sehr wenig zusätzliches Wissen über die Bedienung der Oberfläche und kann das System sehr effizient

Mapping

nutzen. Um dies zu erreichen, ist es sinnvoll auf universellen mentalen Modellen aufzubauen. Solche universellen Modelle sind sehr allgemein, lassen sich aber an vielen verschiedenen Stellen wiederfinden:

**Beispiel: Universelle mentale Modelle**

- Regler weiter auf führt zu größerem Effekt
- Knopf drücken löst eine Aktion aus
- Hyperlinks sind unterstrichen

Der Einsatz solcher universellen Modelle führt dazu, dass bestimmte Aktivitäten fest in den mentalen Modellen der Nutzer verankert sind. Im Allgemeinen bietet das den Vorteil, dass der Benutzer keinen zusätzlichen Lernaufwand für die Bedienung der UI hat, kann aber auch zur Verwirrung führen (s. Aufgabe 23).

Die Grundlage beim Einsatz von universellen mentalen Modellen bilden sogenannte Image Schemata:

**Definition: Image Schemata**

Image Schemata sind grundlegende Bausteine wiederkehrender basaler Erfahrungen.

Image Schemata entstehen automatisch durch Lebenserfahrung. Bei der Interaktion mit der Umwelt (und auch mit Computern) werden ähnliche, wiederkehrende dynamische Wahrnehmungs- und Verhaltensmuster abstrahiert und multimodal als Image Schemata repräsentiert. In den Image Schemata ist zum Beispiel gespeichert, welche Richtung „mehr“ und welche Richtung „weniger“ bedeutet. Das kann dann im Interaktionsdesign bei Slidern und anderen Reglern verwendet werden.

Image Schemata sind abhängig von der Kultur, in der ein Individuum aufwächst. Bei Reglern fällt dies insbesondere bei Sprachen auf, in denen die Leserichtung anders herum ist. Hier sind also die entsprechenden Image Schemata anders, was beim Einsatz der (universellen) mentalen Modelle berücksichtigt werden muss.

## 8.4 Design-Regelwerke

Beim Interaktionsdesign kommen alle Aspekte aus den vorangegangenen Abschnitten zusammen. Man muss gute Mappings einsetzen und ein passendes konzeptionelles Modell entwickeln, das die mentalen Modelle der Nutzer und Affordanzen berücksichtigt. Für die meisten Fälle gibt es keine expliziten Vorgaben, wie *gutes* Design auszusehen hat. Allerdings können Design-Regelwerke auf verschiedenen Ebenen helfen. Die verschiedenen Ebenen von Design-Regelwerken sind in Tabelle 8.2 in absteigender Konkretisierung dargestellt.

Regelwerk	Beispiele
<b>Guidelines</b>	iOS Human Interface Guidelines
<b>Standards</b>	EN ISO 9241
<b>Patterns</b>	Wizards, Grid Layouts
<b>Heuristiken</b>	Shneidermanns 8 Goldene Regeln
<b>Prinzipien</b>	Goldener Schnitt, Gestalttheorie

Tabelle 8.2: Design-Regelwerke

### 8.4.1 Prinzipien

Prinzipien sind sehr abstrakte Regeln für die Gestaltung. In der Regel sind diese Prinzipien unabhängig von der verwendeten Technologie und dem Kontext, in dem sie angewendet werden. Gestaltungsprinzipien können aus Erkenntnissen der Grundlagenwissenschaften wie Psychologie oder den Kognitionswissenschaften abgeleitet werden und bilden ein wichtiges Fundament für das Interaktionsdesign.

#### Goldener Schnitt

Ein Beispiel für ein solches Gestaltungsprinzip ist der Goldene Schnitt. Dieser besagt:

##### Definition: Goldener Schnitt

Teilt man die Strecke  $c$  in die Teilstrecken  $a$  und  $b$  nach dem goldenen Schnitt, so muss gelten:

$$\Phi = \frac{a}{b} = \frac{b}{c} = \frac{b}{a+b} = \frac{1}{1.61803\dots}$$

Vom Menschen wird das Teilungsverhältnis  $\Phi$  als besonders ästhetisch wahrgenommen. Daher kann man ihn auch an vielen Stellen in der Kultur, Architektur, Malerei usw. finden. Beim Interaktionsdesign kann man den goldenen Schnitt z.B. einsetzen, um die Aufteilung eines Fensters in zwei Bereiche ästhetisch ansprechend zu gestalten.

Der Ursprung des Goldenen Schnitts liegt in der Natur. Dort findet man das Verhältnis  $\Phi$  an vielen Stellen wieder. Insbesondere entspricht dies den Proportionen des menschlichen Körpers (z.B. Position der Augen im Gesicht).

#### Gestaltgesetze

Die Gestaltgesetze sind einige der wichtigsten Prinzipien der Gestaltung. Die Gestaltgesetze befassen sich damit, wie der Mensch Formen wahrnimmt und bilden eine wichtige Grundlage für konkretere Design-Regelwerke. Die Gestaltgesetze lassen sich in vier Grundprinzipien zusammenfassen:

**Multistabilität** Tendenz zur mehrdeutigen Wahrnehmung

**Invarianz** Erkennen einfacher geometrischer Objekte, unabhängig von Translation, Ausrichtung, Scherung, ...

**Emergenz** Erkennen von komplexen Objekten aus einfachen Mustern

**Vergegenständlichung** Erkennen von räumlichen Strukturen ohne eigentliche Präsentation

Mit den vier Grundprinzipien werden sehr viele Aspekte und Kriterien zusammengefasst, die man im Interaktionsdesign anwenden kann. Wir werden davon hier nur einige genauer betrachten:

**Figur & Hintergrund** Ein wichtiger Aspekt der Objektwahrnehmung ist, dass Menschen eine Tendenz haben, in Bildern einen Vorder- und einen Hintergrund zu erkennen. Das bekannteste Beispiel für dieses Phänomen ist in Abbildung 8.2 dargestellt. Hier kann man entweder eine schwarze Vase auf weißem Hintergrund oder zwei weiße Gesichter auf einem schwarzen Hintergrund erkennen. Das Gehirn ist normalerweise nicht in der Lage, beide Dinge gleichzeitig zu erkennen, sondern muss zwischen den beiden Möglichkeiten hin- und herschalten.

**Binnengliederung** Mit Binnengliederung werden verschiedene Kriterien bezeichnet, die dafür sorgen, dass das Gehirn mehrere Objekte als zusammengehörig wahrnimmt. Diese Kriterien werden bei nahezu allen WIMP-GUIs eingesetzt, um die Benutzeroberfläche zu strukturieren.

**Geschlossenheit** Der Mensch hat eine starke Tendenz, geschlossene Objekte wahrzunehmen. Diese Tendenz kann man im Alltag beobachten, wenn man ein Objekt betrachtet, das teilweise verdeckt ist. Das Gehirn ergänzt dann automatisch den fehlenden Teil. Durch abgeschlossene Objekte entsteht außerdem eine natürliche Trennung. Das kann man im Interaktionsdesign nutzen, indem man beispielsweise zusammengehörige Optionen durch einen Rahmen von anderen Optionen trennt.

**Ähnlichkeit** Ähnliche Elemente werden vom Gehirn ebenfalls als Gruppe wahrgenommen. Kriterien für Ähnlichkeit können z.B. Form, Farbe oder Bewegung sein.

**Nähe** Objekte, die nahe beieinander sind, werden ebenfalls als zusammengehörig erkannt. Dies ist ein sehr dominantes Kriterium, das oft bei Form-Fill Interfaces eingesetzt wird. Hier sind die Labels für die Datenfelder sehr nahe bei den Datenfeldern, sodass der Nutzer diese Zugehörigkeit erkennen kann.

Es ist sehr selten, dass bei einem Interface nur ein Gestaltgesetz eingesetzt wird. Normalerweise ist das auch gar nicht möglich. Im Interaktionsdesign setzt man verschiedene Gestaltgesetze gemeinsam ein, um eine bestimmte Wirkung zu erzielen. Dabei kann es sowohl vorkommen, dass sich zwei Gesetze gegenseitig unterstützen als auch dass sich

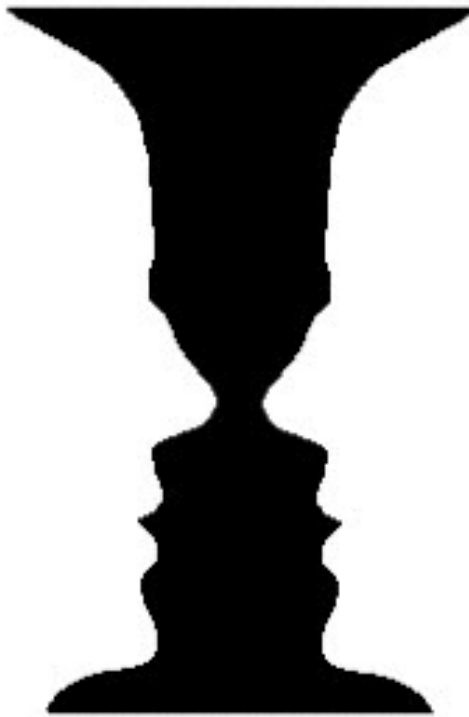


Abbildung 8.2: Vase oder Gesicht?

zwei Gesetze widersprechen. Wenn es einen Konflikt zwischen zwei Gesetzen gibt, muss das nicht notwendigerweise ein Problem sein. Viel mehr kann man die Möglichkeiten der verschiedenen Gesetze nutzen, um komplexe Interfaces zu entwickeln. So kann man zum Beispiel das Gesetz der Ähnlichkeit nutzen, um verschiedene Gruppen als zusammengehörig zu identifizieren und das Gesetz der Nähe nutzen, um die Elemente innerhalb einer Gruppe von den anderen Gruppen zu trennen.

#### **8.4.2 Heuristiken/Goldene Regeln**

Auf der nächsten Ebene der Design-Regelwerke stehen die Heuristiken. Heuristiken bzw. Goldene Regeln sind spezifische Gestaltungsregeln, die meist einfach, aber konkret formuliert sind und einen Empfehlungscharakter haben.

##### **Shneidermans 8 Goldene Regeln**

Ben Shneiderman war einer der Begründer der Mensch-Computer-Interaktion in den 80er Jahren. Shneidermann hat eine Heuristik mit 8 Regeln aufgestellt, die dem Interaktionsdesigner helfen sollen, eine gute grafische Benutzerschnittstelle zu designen:

- 1. Konsistenz**

Verwende Styleguides und Konventionen

2. **Berücksichtige unterschiedliche Erfahrungen**

Die UI sollte möglichst jedem eine passende Interaktionsform anbieten, z.B. Menüs für Anfänger und Tastenkürzel für erfahrene Benutzer.

3. **Rückmeldungen auf Benutzeraktionen**

Sobald eine Aktion bei der Software angekommen ist, sollte der Benutzer darüber informiert werden. Das ist insbesondere wichtig, wenn die Ausführung der Aktion nicht sofort ein Ergebnis liefert.

4. **Abgeschlossene Operationen**

Die Schritte einer mehrstufigen Operation sollten im Zusammenhang dargestellt werden.

5. **Fehler verhindern**

Benutzerfehler sollten verhindert werden, indem Auswahlalternativen angeboten werden und die Darstellung eindeutig ist.

6. **Einfache Rücksetzmöglichkeiten**

Wenn der Benutzer die Möglichkeit hat, Aktionen rückgängig zu machen, entsteht ein Sicherheitsgefühl. Dadurch kann der Benutzer explorativ das System kennen lernen.

7. **Benutzer bestimmt Kontrollfluss**

Der Benutzer muss das Gefühl haben, die Anwendung steuern und kontrollieren zu können.

8. **Geringe Belastung des Arbeitsgedächtnis**

Der Benutzer sollte möglichst Erkennen anstatt sich erinnern zu müssen. Menüs sollten besser breit als tief aufgebaut sein.

### 8.4.3 Patterns

Patterns sind Lösungen eines grundlegenden Gestaltungsproblems innerhalb einer bestimmten Domäne. Ein Pattern ist somit eine Art Muster oder Vorlage, die für die Lösung eines bestimmten Problems eingesetzt werden kann. Wenn man ein etabliertes Pattern zur Lösung eines Problems einsetzt, hat dies für den Benutzer den Vorteil, dass dieser bereits ein mentales Modell von dem Pattern hat und dadurch sofort weiß, wie er mit dem System umgehen muss.

Es gibt sehr viele Patterns. Wir betrachten hier beispielhaft nur zwei:

**Feature/Search/Browse** Dieses Pattern besteht aus drei Elementen:

1. Artikel oder Produkte (*engl. Feature*)
2. Suchfunktion (*engl. Search*)



3. Liste von Items, die durchstöbert werden können (*engl. Browse*)

Insbesondere bei E-Commerce Systemen wie beispielsweise Amazon wird dieses Pattern eingesetzt. Im Allgemeinen lässt sich das Pattern aber immer dann gut anwenden, wenn der Benutzer aus einer sehr großen Menge von Objekten wählen muss.

**Breadcrumbs** Breadcrumbs sind ein Weg, dem Benutzer mitzuteilen, an welcher Stelle in einer Hierarchie er sich befindet. Normalerweise bietet dieses Pattern durch Klicken auf eines der Elemente die Möglichkeit, in der Hierarchie zurückzuspringen.

#### 8.4.4 Standards

Standards sind spezifische Gestaltungsregeln mit verbindlichem Charakter. Das nicht Einhalten von Standards kann somit auch rechtliche Konsequenzen mit sich führen. Standards haben nur einen geringen Anspruch auf Allgemeingültigkeit und sind daher nicht für alle Anwendungsbereiche geeignet. Im Bereich Interaktionsdesign ist beispielsweise die Norm DIN EN ISO 9241 - „Ergonomie der Mensch-System-Interaktion“ besonders relevant. Diese Norm ist in viele Abschnitte unterteilt, von denen hier zwei aufgelistet sind:

**DIN EN ISO 9241-11: Gebrauchstauglichkeit** Dieser Teil der Norm besagt, dass Software mit einer guten Gebrauchstauglichkeit (*engl. Usability*) bestimmte Leitkriterien erfüllen muss:

1. **Effektivität** zur Lösung einer Aufgabe
2. **Effizienz** der Handhabung des Systems
3. **Zufriedenheit** der Nutzer einer Software

Die Quantisierung dieser Kriterien erfolgt zum Beispiel über Fragebögen, biometrische Daten oder Messungen der Zeit bzw. Fehler beim Erfüllen einer Aufgabe.

#### DIN EN ISO 9241-307: Pixelfehlerklassen

#### 8.4.5 Guidelines

Guidelines sind Regelwerke, die einen *a priori* verbindlichen Charakter von Standards haben. Guidelines werden geschrieben, um interne (innerhalb eines einzelnen Programms) und externe (über mehrere Programme hinweg) Konsistenz einer Umgebung sicherzustellen. Die meisten Betriebssysteme (z.B. Windows, iOS, Android) haben solche Guidelines, in denen vorgegeben wird, welche Design-Aspekte bei Programmen für dieses System berücksichtigt werden müssen.

## Beispiel

- OS X Human Interface Guidelines
- Windows Desktop Application Design
- iOS Human Interface Guidelines
- Android Material Design

# 9 Evaluierung

Evaluierung ist ein wesentlicher Bestandteil der Mensch-Computer-Interaktion. Es geht darum, Informationen über die Perfomanz sowie über Probleme und Erfahrungen der Benutzer beim Umgang mit interaktiven Systemen zu sammeln, um Probleme im System zu identifizieren und um verschiedene Systeme miteinander vergleichen zu können.

## 9.1 Arten der Evaluierung

### 9.1.1 Warum?

Evaluierung wird in Form von Studien durchgeführt und in der Mensch-Computer-Interaktion eingesetzt, wenn ein neues System oder eine neue UI entwickelt wird. Es gibt verschiedene Gründe, aus denen man eine Evaluierung vornehmen möchte:

**Exploration** Die Exploration findet in einem sehr frühen Stadium des Entwicklungsprozesses statt, um eine qualitative und informelle Erkundung der Benutzeranforderungen zu liefern. Dies findet meistens in Form von Umfragen oder Fokusgruppen statt, in denen man sich überlegt, welche Anforderungen das neue System erfüllen muss.

**Beurteilung** Eine Beurteilung findet statt, um den aktuellen Stand der UI-Entwicklung einzuschätzen und Alternativen zu vergleichen. Eine solche Evaluierung kann eingesetzt werden, um sich zwischen zwei oder mehr Designalternativen zu entscheiden.

**Vergleich** Beim Vergleich werden verschiedene, konkrete UIs miteinander verglichen und nach ausgewählten Kriterien beurteilt (z.B. ist das neue System schneller als das alte?).

**Validierung** Am Ende der Entwicklungsphase werden Behauptungen und Hypothesen überprüft, die man am Anfang aufgestellt hat. Solche Behauptungen können sich zum Beispiel auf die Effizienz oder Genauigkeit einer entwickelten UI beziehen.

### 9.1.2 Wann?

Aus den obigen Evaluierungstypen ergeben sich verschiedene Zeitpunkte, zu denen Evaluierung durchgeführt wird:

**Vor der Entwicklung** findet Evaluierung statt, um eine **Anforderungsanalyse** zu erstellen. Dazu spricht man mit den Nutzern und identifiziert dabei Probleme bei der Bedienung aktueller Systeme. Mithilfe dieser Daten kann dann ein neues System entwickelt werden.

**Während der Entwicklung** gibt es oft verschiedene Designalternativen (z.B. miteinander konkurrierende Gestaltgesetze) zwischen denen man sich entscheiden muss. Um diese Entscheidung möglichst benutzerfreundlich zu machen, kann man bereits während der Entwicklung Evaluierungen durchführen, um verschiedene Alternativen zu testen. Das Resultat ist normalerweise eine Verfeinerung der Anforderungsanalyse, die schon vor der Entwicklung erstellt wurde.

**Nach der Entwicklung** kann man durch Evaluierung prüfen, ob die Anforderungen an das neue System erfüllt wurden. Dies kann durch eine Validierung oder einen Vergleich mit anderen Systemen erfolgen.

Bei Evaluierungen, die nach der eigentlichen Produktentwicklung stattfinden, spricht man auch von summativer Evaluierung. Vor und während der Entwicklung spricht man von formativer Evaluierung, da diese den Entwicklungsprozess beeinflussen und das entwickelte Produkt *formen* kann.

### 9.1.3 Feld- und Laborstudien

Eine weitere Eigenschaft, die Evaluierungen voneinander unterscheidet, ist der Ort der Durchführung. Man unterscheidet hier hauptsächlich Feld- und Laborstudien:

#### Definition: Feldstudie

Eine Feldstudie ist eine systematische, wissenschaftliche Beobachtung unter natürlichen Bedingungen.

Feldstudien zeichnen sich besonders dadurch aus, dass es viele Faktoren gibt, die nicht durch die Experimentatoren beeinflusst werden können. Obwohl dies die Evaluierung erschwert, ist eine Feldstudie nötig, um die Tauglichkeit eines Produktes in der echten Welt zu beurteilen. Insbesondere Ablenkung, Lärm und Mitmenschen sind unkontrollierbare Faktoren, die bei einer Feldstudie eine wichtige Rolle spielen.

#### Definition: Laborstudie

Eine Laborstudie ist ein wissenschaftliches Mittel, um mit Hilfe von Laborexperimenten eine bestimmte Arbeitshypothese zu testen.

Im Gegensatz zu Feldstudien sind bei Laborstudien nahezu alle Einflussfaktoren kontrolliert. Kontrollierte Faktoren erstrecken sich über das getestete Produkt hinaus und umfassen z.B. die Farbe der Wände, Lichtverhältnisse und Einflüsse durch Mitmenschen. Der Nachteil von Laborstudien ist allerdings, dass man nur schwer Rückschlüsse auf den Einsatz in der echten Welt ziehen kann.

Normalerweise bevorzugen Wissenschaftler Labrostudien, da man hier sehr genau kontrollieren kann, welche Einflüsse sich auf das getestete Produkt auswirken. Das Problem

ist aber, dass in Laborstudien nie alle Einflussfaktoren der echten Welt realistisch nachgebildet werden können. In den letzten Jahren hat sich daher ein Trend herausgebildet, Feld- und Laborstudien in gewissem Maße zu kombinieren, indem man die natürliche Umgebung des Benutzers im Labor abbildet. Die Benutzer müssen also wie bei einer Laborstudie ins Labor gehen, in dem der Benutzer weiterhin durch Kameras und Menschen beobachtet wird, allerdings sieht der Raum, in dem das Experiment stattfindet, beispielsweise aus wie ein Wohnzimmer.

#### 9.1.4 Ergebnisse

Ergebnisse von Evaluierungen können unterschiedliche Formen haben:

1. Schwer quantifizierbare Beschreibungen (z.B. Lösungsansätze)
2. Benutzeraussagen (z.B. Präferenzen oder Selbsteinschätzungen)
3. Quantitativ messbare Ergebnisse
4. Technisch Aufzeichnbar und auswertbar

Wir unterscheiden die Ergebnisse abhängig davon, ob diese Subjektiv (1. und 2.) oder objektiv (3. und 4.) sind. Für die Auswertung der Ergebnisse muss man zusätzlich noch zwischen qualitativen und quantitativen Daten unterscheiden:

##### Definition: Quantitative Ergebnisse

*„When, among a set of observations, any single observation is a number that represents an amount or a count, then the data are quantitative.“*

— Witte & Witte: Statistics, 2009

Als Quantitative Evaluierung bezeichnet man Evaluierungen, die quantitative Ergebnisse liefern, also Ergebnisse, die sinnvoll durch Zahlen ausgedrückt werden können (z.B. Anzahl der Fehler, Zeit, Anzahl der Tastendrucke, ...)

##### Definition: Qualitative Ergebnisse

*„When, among a set of observations, any single observation is a word, or a sentence, or a description, or a code that represents a category then the data is qualitative.“*

— Witte & Witte: Statistics, 2009

Analog bezeichnet man als Qualitative Evaluierung solche Evaluierungen, die qualitative Ergebnisse liefern, also Ergebnisse, die sich nicht sinnvoll durch Zahlen beschreiben lassen (z.B. Interviews).

Qualitative und quantitative Ergebnisse ergänzen sich häufig sinnvoll. Während sich quantitative Daten sehr gut für statistische Auswertungen eignen und belastbare Ergebnisse produzieren, runden erst qualitative Daten die Ergebnisse einer Studie ab.

## 9.2 Analytische Methoden

Evaluierung kann durch analytische und empirische Untersuchungen durchgeführt werden. In diesem Abschnitt betrachten wir die analytischen Methoden genauer.

Analytische Methoden analysieren ein System durch reine Analyse. Sie liefern dadurch Erklärungen von Arbeitsweisen, Bestandteilen oder Eigenschaften des Systems oder eines Produktes. In der Regel werden dabei keine Testpersonen mit einbezogen.

### Beispiel: Analytische Methoden

- KLM-GOMS
- Cognitive Walkthrough
- Heuristische Evaluation auf Basis von Nielsen oder Shneiderman

Analytische Untersuchungen werden im Normalfall durch Expertengruppen vorgenommen. Diese beurteilen ein System nach klaren Kriterien (z.B. GOMS oder heuristische Verfahren). Normalerweise reicht es, wenn 3-5 Experten die Beurteilung vornehmen. Die Ergebnisse von solchen Untersuchungen werden dann mithilfe von Bewertungsskalen festgehalten. Das Ziel von Bewertungsskalen ist, die Meinungsbilder der Experten in quantitative Form zu bringen. Ein häufig eingesetztes Verfahren ist die Likert-Skala:

Bei der Likert-Skala ordnet man einer Anzahl von Bewertungen Kästchen zum Ankreuzen zu. Die Optionen sind dabei häufig „Ich stimme voll zu“, „Ich stimme ein wenig zu“... Meistens haben solche Skalen eine ungerade Anzahl von Optionen, damit es eine klare Mitte gibt. Man kann aber auch genau darauf verzichten eine Mitte anzubieten, indem man eine gerade Anzahl von Optionen anbietet.

## 9.3 Empirische Methoden

Empirische Methoden bezeichnen alle Formen von Evaluation, die durch Messung oder anderweitige Sammlung in Experimente, Beobachtungen oder Befragungen Daten erheben, auf deren Basis wissenschaftliche Aussagen gemacht werden können. Im Gegensatz zu analytischen Methoden beziehen diese Verfahren in der Regel Testpersonen ein. Die Ergebnisse, die man durch empirische Untersuchungen erhält, eignen sich meistens sehr gut, um Aussagen über die Qualität und Quantität der Bedienung eines interaktiven Systems machen zu können, nicht jedoch, um dies zu begründen. Dafür setzt man analytische Verfahren ein.

Für erfolgreiche empirische Untersuchungen müssen vier Eigenschaften erfüllt sein:

**Objektivität** bedeutet, dass die erhobenen Daten unabhängig von der Messmethode, Erwartung und Hypothesen des Experimentators sind.

**Reproduzierbarkeit** bedeutet, dass das Experiment hinreichend genau beschrieben ist, so dass es wiederholbar wird.

**Validität** bedeutet, dass die Ergebnisse nur das messen, was sie messen sollen (interne Validität) und repräsentativ für die Allgemeinheit sind (externe Validität).

**Relevanz** bedeutet, dass die Ergebnisse tatsächlich neue Erkenntnisse liefern.

Damit alle diese Eigenschaften erfüllt sind, versucht man möglichst empirische Untersuchungen in Form von kontrollierten Experimenten durchzuführen. Kontrollierte Experimente sind empirische Studien, bei denen alle relevanten Einflussfaktoren kontrolliert werden können. Auf diese Weise soll mit möglichst wenig Versuchen die Kausalität zwischen Einflussfaktoren und Zielgrößen ermittelt werden.

Kontrolliertes  
Experiment

Zielgrößen  
Einflussfaktoren

Möchte man eine empirische Untersuchung durchführen, legt man zuerst die Variablen fest. Wir unterscheiden verschiedene Variablentypen:

**Unabhängige Variablen** sind Variablen, die im Laufe des Experimentes kontrolliert verändert werden können und sich sonst nicht ändern (z.B. Eingabegerät oder Interface-Design).

**Abhängige Variablen** sind Variablen, die sich durch das Anpassen von unabhängigen Variablen ändern (z.B. Zeit oder Anzahl der Fehler).

**Kontrollvariablen** könnten einen Einfluss auf abhängige Variablen haben und sollten daher einmal erhoben und sonst konstant gehalten werden (z.B. Beleuchtung, Temperatur).

**Zufallsvariablen** könnten sich auf abhängige Variablen auswirken, bleiben aber dem Zufall überlassen. Dass es Zufallsvariablen gibt ist für die Generalisierbarkeit (externe Validität) wichtig. Meistens haben Zufallsvariablen etwas mit den Eigenschaften der Testpersonen zu tun (z.B. Größe oder Gewicht).

**Störfaktoren** sind Variablen, die sich mit einer unabhängigen Variable verändern. Sie sind problematisch, da nun unklar ist, ob die Kausalität zwischen der unabhängigen Variable oder dem Störfaktor und der abhängigen Variable besteht (z.B. unterschiedliche Maustechnologien).

Nachdem man die Variablen festgelegt hat, kann man das Experiment durchführen. Bei der Durchführung versucht man kontrolliert die unabhängigen Variablen zu verändern und misst die Auswirkungen auf die abhängigen Variablen. Dabei sollte man versuchen unerwünschte Nebeneffekte wie Ermüdung oder Lerneffekte zu verhindern, indem man zum Beispiel eine Gruppe erst Alternative A und eine andere Gruppe erst Alternative B testen lässt.

## 9.4 Datenanalyse

Die Datenanalyse ist der Prozess erhobene Daten so aufzubereiten, dass man daraus Schlussfolgerungen ziehen kann. Wir beschäftigen uns hier hauptsächlich mit der Analyse von quantitativen Daten. Qualitative Daten zu analysieren ist deutlich komplexer. Im

einfachsten Fall kann man qualitative Daten einfach durch einen Transformationsprozess in quantitative Daten umwandeln (z.B. indem man das Vorkommen bestimmter Wörter zählt).

### 9.4.1 Deskriptive Statistik

Eine Deskriptive Statistik beschreibt die Ergebnisse eines Experimentes und ist somit ein Resultat der Datenanalyse. Eine Deskriptive Statistik enthält normalerweise:

- Maximalwert und Minimalwert
- Mittelwert (Durchschnitt)
- Standardabweichung (Durchschnittliche Abweichung vom Mittelwert)
- Median (Mittlerer Wert in geordneter Liste)
- Modus (Am häufigsten auftretender Wert)

Wird eine deskriptive Statistik veröffentlicht, kann dies ebenfalls verschiedene Formen annehmen. Meistens werden Text, Grafiken und Diagramme verwendet.

#### Aufgabe 24: Deskriptive Statistik

Geben Sie die relevanten Daten für eine deskriptive Statistik zu den Datenwerten

$$\{0, 1, 1, 2, 3, 4, 4, 4, 5\}$$

an.

### 9.4.2 Subgruppenanalyse

Eine andere Art der Datenanalyse ist die Subgruppen- bzw. Clusteranalyse. Hier teilt man die Probanden nach bestimmten eigenschaften in Subgruppen auf und analysiert dann, welche unabhängigen Variablen zu dieser Aufteilung geführt haben. Am Ende kann man daraus ableiten, ob die Änderungen im Experiment signifikant waren (d.h. ob die Wahrscheinlichkeit hoch ist, dass die gemessenen Ergebnisse bei anderen Versuchsgruppen ebenfalls auftreten). Man unterscheidet bei der Subgruppenanalyse:

**Subgruppenanalyse** Hier sind die Subgruppen schon vor der Durchführung des Experimentes bekannt. Die Probanden werden zum Beispiel nach unterschiedlichen Konditionen aufgeteilt.

**Clusteranalyse** Hier werden die Subgruppen auf Basis der gemessenen Daten eingeteilt.

Normalerweise ist eine Subgruppenanalyse nur ein Teil der kompletten Datenanalyse. Beim Veröffentlichen der Daten wird dann wieder eine deskriptive Statistik verwendet.



### 9.4.3 Grafiken

Grafiken sind ein beliebtes Mittel zur Veranschaulichung von Datenanalysen. Allerdings geht aus Grafiken normalerweise nicht hervor, ob die Resultate statistisch signifikant sind (dazu muss man Signifikanztests durchführen). Daher kann man Diagramme nur als unterstützendes Material einsetzen, das die Ergebnisse veranschaulicht. Ein Diagramm kann keine deskriptive Statistik ersetzen.

## 9.5 Statistische Tests

Statistische Tests werden nach der Durchführung eines quantitativen Experiments eingesetzt, um festzustellen, ob die Ergebnisse signifikant sind oder nicht. Statistische Tests haben sehr genaue formale Vorgaben, die man einhalten muss. Diese beziehen sich auf folgende Eigenschaften eines Experiments:

**Hypothesen** Jedes Experiment hat mindestens zwei Hypothesen:

**Hypothese** Die Hypothese ist der vorhergesagte Einfluss der unabhängigen Variablen auf die abhängigen Variablen.

**Nullhypothese** Die Nullhypothese besagt, dass es keinen Einfluss der Konditionen auf die Messwerte gibt.

Das Ziel eines statistischen Tests ist es, die Hypothese zu belegen und die Nullhypothese zu verwerfen.

**Variablentypen** In einem Experiment gibt es verschiedene Arten von quantitativen Variablen:

**Nominal** Nominale Variablen sind Merkmale, die keine natürliche Ordnung haben (z.B. das Geschlecht).

**Ordinal** Ordinale Variablen sind Merkmale mit natürlicher Ordnung (z.B. „jeden Tag“, „einmal in der Woche“).

**Intervall** Intervallvariablen sind Merkmale mit natürlicher Ordnung und der Zusatzbedingung, dass alle Werte den gleichen Abstand voneinander haben. Es gibt keinen absoluten Nullpunkt (z.B. Likert-Skalen mit als gleich ansehbaren Abständen).

**Ratio** Ratio-Variablen sind Merkmale mit natürlicher Ordnung, gleichen Abständen zwischen den Werten und mit absolutem Nullpunkt (z.B. Alter oder Einkommen).

**Messwiederholungen** Man muss für statistische Tests unterscheiden, wie ein Experiment mit Messwiederholungen umgeht:

**Within-Subjects Design** Die Probanden haben alle verfügbaren Konditionen durchgemacht.

**Between-Subjects Design** Jeder Proband hat nur einen Teil der Konditionen durchgemacht.

Nachdem alle Eigenschaften des Experiments klar sind, kann man den konkreten statistischen Test aus Tabelle 9.1 auswählen.

	Within-/Between-Subjects Design	Typ der abhängigen Variable		
		Intervall/Ratio (Normalität angenommen)	Intervall/Ratio (Normalität nicht angenommen), Ordinal	Dichotomy (Binomial)
Mittelwertsvergleich von zwei Gruppen	Between	Unpaired T-Test	Mann-Whitney Test	Fisher's Test
	Within	Paired T-Test	Wicoxon Test	McNemar's Test
Mittelwertsvergleich von mehr als zwei Gruppen	Between	ANOVA	Kruskal-Wallis Test	Chi-Square Test
	Within	Repeated-measures ANOVA	Friedman Test	Cochran's Q-Test
Zusammenhang zwischen zwei Variablen finden	Beide	Pearson Correlation	Spearman Correlation	Cramer's V
Wert vorhersagen mit einer unabhängigen Variable		Linear/Non-linear Regression	Non-parametric Regression	Logistic Regression
Wert vorhersagen mit mehreren unabhängigen Variablen oder binomialen Variablen		Multiple linear/non-linear Regression		Multiple Logistic Regression

Tabelle 9.1: Statistische Tests

## 9.6 DECIDE-Framework

Bevor ein Experiment evaluiert werden kann, müssen einige W-Fragen beantwortet werden: Wer, Warum, Was, Wie, Wann und Wo. DECIDE bietet ein Framework, welches dabei hilft, Benutzerstudien zu planen und durchzuführen und dient als Checkliste für die Evaluierung. Es setzt sich zusammen aus:

### D etermine Goals

Zuerst werden die Ziele der Evaluation festgelegt. Das beinhaltet die Fragen

- Wer will die Evaluierung?
- Warum soll die Evaluierung durchgeführt werden?
- Was soll herausgefunden werden?

### E xplore the Question

Als nächstes wird die Frage aufgestellt, die durch die Evaluierung beantwortet werden soll. Größere Fragen müssen oft in kleinere Fragen zerlegt werden, die dann in mehreren kleinen Evaluierungen zerlegt werden.

- **C**hoose Evaluation Method

In diesem Schritt wird die Methode der Evaluierung in Abhängigkeit von den Zielen und der Fragestellung ausgewählt. Meistens benutzt man nicht nur eine Methode, sondern mehrere, um sich der Fragestellung von verschiedenen Blickwinkeln aus zu nähern.

- **Identify Practical Issues**  
Hier wird festgelegt, welche Aspekte den reibungslosen Ablauf der Evaluierung beeinflusst. Das beinhaltet Laboraustattung, Zeit und Budget und die Expertise. Damit dieser Schritt geklärt werden kann, führt man meistens Pilot-Studien durch, in denen die Studie probeweise an sich selbst oder Freunden und Bekannten durchgeführt wird. Bei solchen Pilot-Studien kann man meistens sehr gut sehen, wenn es irgendwelche Probleme bei der Durchführung gibt.
- **Decide on Ethical Issues**  
Vor der eigentlichen Evaluierung müssen ethische Fragestellungen geklärt werden. Dafür ist die Zertifizierung einer Ethikkommission notwendig (z.B. dem Institutional Review Board).
- **Evaluate**  
Nachdem alle anderen Schritte erfolgreich abgeschlossen wurden, kann die Evaluation durchgeführt werden. Nach der Evaluation müssen die Daten ausgewertet werden (z.B. durch statistische Tests).

# Lösungen zu den Aufgaben

## Aufgabe 1

Der Kontext im PACT-Framework ist ein Umfeld, in dem Technologie zum Einsatz kommt. Das kann z.B. eine Firma oder die ganze Bevölkerung eines Landes sein. In diesem Kontext gibt es bestimmte Aktivitäten, die normalerweise von Menschen ausgeführt werden. Die Technologie ermöglicht nun, dass die Aktivitäten einfacher oder schneller ausgeführt werden können. In diesem Sinne stellt der Kontext gewisse **Anforderungen** an die Technologie, um die Aktivitäten zu optimieren.

Auf der anderen Seite kann die Technologie nur so viel, wie der aktuelle Forschungsstand erlaubt. Die **Möglichkeiten** der Technologie und der damit verbundenen Optimierung von Aktivitäten sind also durch die Fortschritte in der Forschung eingeschränkt. So entsteht eine Art Sog vom Kontext zur Technologie mit dem Ziel die Aktivitäten zu optimieren und damit das Leben für die Menschen zu vereinfachen.

## Aufgabe 2

Damit eine Bewegung flüssig wahrgenommen wird, müssen mehr Bilder pro Sekunde gezeigt werden, als der perzeptuelle Prozessor verarbeiten kann. Andernfalls erkennt der perzeptuelle Prozessor die einzelnen Bilder. Der kognitive und motorische Prozessor sind für das Erkennen von Bewegungen erstmal irrelevant, da weder die Bewegung interpretiert werden muss, noch eine Aktion ausgeführt wird. Für den perzeptuellen Prozessor gilt:

$$\tau_P = 100[50 \sim 200]\text{ms} = 0.1[0.05 \sim 0.2]\text{s}$$

Damit mehr Bilder pro Sekunde (*engl. Frames per Second, fps*) gezeigt werden, als der perzeptuelle Prozessor wahrnehmen kann, muss also gelten:

$$\text{fps} > \frac{1}{\tau_P} = \frac{1}{0.05} = 20$$

Man beachte, dass hier die *kürzeste* Zeit des perzeptuellen Prozessors verwendet werden muss, damit alle Menschen die Bewegung flüssig sehen. Also muss es mindestens 20 Bilder pro Sekunde geben, damit alle Menschen eine Bewegung flüssig sehen können.

*Da es sich bei den Laufzeiten für die Prozessoren um empirische Daten handelt, gibt es bei 20 Bildern noch keine Garantie, dass alle Menschen die Bewegung tatsächlich flüssig sehen. Aktuelle Film- und Fernsehstandards haben meistens Bildraten, die etwas höher sind als 20 Bilder pro Sekunde.*

### Aufgabe 3

Die (präattentive) Erkennung der roten Dreien benötigt etwa 200 ms. Dazu kommt nun noch die Zeit, die es dauert, um die Dreien zu zählen. Das Zählen selbst ist dabei eine attentive Aufgabe. Es gibt 14 Dreien im Bild. Für jede Drei muss mindestens der kognitive und der motorische Prozessor durchlaufen werden (Durchführen der Addition und Weiterbewegen der Augen). Damit kann man die Bearbeitungszeit auf  $200 \text{ ms} + 14 \cdot 200 \text{ ms} = 3 \text{ s}$  abschätzen. Diese Schätzung ist deutlich realistischer als die ursprünglichen 23,4 s.

### Aufgabe 4

In der Grafik ist die Option *Quit App* farblich hervorgehoben. Farbe ist ein präattentives Merkmal, daher kann diese Option besonders schnell gefunden werden.

### Aufgabe 5

Die umgangssprachliche Wahrnehmung besteht aus drei Schritten:

1. Aufnehmen des Reizes durch das perzeptuelle System.
2. Sensorische Integration (Kombinieren von Reizen zu einem Eindruck).
3. Interpretation und probabilistische Ergänzung zu einem Gesamteindruck.

Hört man eine Sirene in einem Flugzeug, gibt es (etwas vereinfacht) zwei relevante Reize: Einen visuellen Reiz und einen auditiven Reiz. Visuell sehen Sie, dass Sie in einem Flugzeug sind, auditiv hören Sie die Sirene. Bei der Sensorischen Integration werden diese beiden Reize dann kombiniert. Hier gibt es zwei Möglichkeiten:

1. Sie sehen direkt die Quelle der Sirene. In dem Fall gibt es nach diesem Schritt nur noch die kombinierte Information über die Quelle und das Ereignis selbst. In dem Fall ist hier schon klar, dass es sich nicht um ein Polizeiauto handelt.
2. Sie wissen nicht, woher die Sirene kommt.

In beiden Fällen wird als nächstes die Interpretation ausgeführt. Spätestens dort wird die Kombination *Flugzeug* und *Polizeiauto* verworfen, da es sehr unwahrscheinlich ist, ein Auto in einem Flugzeug zu finden.

### Aufgabe 6

Es gilt die Formel:

$$c = \lambda \cdot f$$

Diese kann man nun nach  $\lambda$  bzw.  $f$  umstellen:

$$\lambda = \frac{c}{f} \quad f = \frac{c}{\lambda}$$

Nun muss man nur die gegebenen Werte einsetzen:

1. Die Farbe von Licht mit 500 nm kann man im elektromagnetischen Spektrum (Abbildung 3.3) ablesen. Das Licht ist grün.

$$f = \frac{c}{\lambda} = \frac{300\,000 \text{ km s}^{-1}}{500 \text{ nm}} = \frac{300 \cdot 10^6 \cancel{\text{m}}}{500 \cdot 10^{-9} \cancel{\text{m}} \text{s}} = 6 \cdot 10^{14} \text{ s}^{-1} = 6 \cdot 10^{14} \text{ Hz}$$

2.  $1,4 \cdot 10^{14} \text{ Hz}$ :

$$\lambda = \frac{c}{f} = \frac{300 \cdot 10^6 \cancel{\text{ms}}}{1,4 \cdot 10^{14} \cancel{\text{s}}} \approx 214,29 \cdot 10^{-8} \text{ m} \approx 2143 \text{ nm}$$

2143 nm ist deutlich größer als die 800 nm, die die obere Schwelle des Lichtes bilden, das der Mensch wahrnehmen kann. Daher fällt das Licht in den infraroten Bereich.

### Aufgabe 7

1. Eine Tomate reflektiert fast nur rotes Licht. Bestrahlt man eine Tomate mit blauem Licht, wird dieses fast vollständig absorbiert und somit fast kein Licht wird reflektiert. Solange es neben der blauen Lichtquelle keine andere Lichtquelle im Raum gibt, erscheint die Tomate also gräulich oder schwarz.
2. Eine Fensterscheibe ist nicht komplett Lichtdurchlässig. Ein kleiner Teil des einfallenden Lichtes wird immer reflektiert. Ist es draußen dunkel, wird deutlich weniger Licht von draußen durch das Fenster gelassen, als von drinnen reflektiert wird. Tagsüber ist dies genau umgekehrt. Das Licht, das von draußen durch das Fenster fällt, überlagert den kleinen Anteil des von drinnen reflektierten Lichtes, sodass man hinausschauen kann.
3. Damit ein Objekt von einem Menschen gesehen wird, müssen reflektierte

Lichtstrahlen das Auge erreichen. Wenn sie das Auge erreicht haben, wird das Licht dort absorbiert. Das heißt aber, dass eben dieses Licht danach kein anderes Auge mehr erreichen kann. Für eine unsichtbare Person gelten dieselben physikalischen Gesetze. Das heißt eine vollständig unsichtbare Person nichts sehen kann, da kein Licht das Auge erreicht. Falls doch etwas Licht das Auge erreicht, können Beobachter eben dieses Licht nicht mehr sehen. Das würde sich durch eine Art Schatten an der Stelle der „unsichtbaren“ Person bemerkbar machen.

### **Aufgabe 8**

Damit man erfolgreich ein Objekt fokussieren kann, sind mehrere Variablen wichtig. Neben der Brennweite spielt auch der Abstand zwischen Linse und Fokusebene (Fokusedistanz) sowie der Abstand zwischen Linse und Bildebene (Bildweite) eine Rolle. Im Auge ist die Bildweite der Abstand zwischen Linse und Retina. Durch die normale genetische Varianz kommt es nun vor, dass nicht alle Menschen gleich große Augen haben. Genauso kann es passieren, dass die Bildweite des Auges etwas zu groß oder zu klein ist und somit die Objekte, die man fokussieren will, unscharf erscheinen.

Eine solche Anomalie des Auges heißt nicht, dass das Auge nun keine Fokusebene mehr hat. Die Fokusebene liegt nur woanders, als sie das Gehirn erwartet. Deshalb können manche Brillenträger mit einiger Konzentration auch ohne Brille scharf sehen, indem Sie bewusst die Aufmerksamkeit von der physikalischen Fokusebene entfernen (ähnlich wie im Beispiel, wenn man seine Aufmerksamkeit über den ausgestreckten Arm auf den unscharfen Hintergrund lenkt).

### **Aufgabe 9**

Es gibt verschiedene Ausprägungen der Farbschwächen. Eine Möglichkeit die Rot/Grün-Schwäche zu erklären, ist die folgende:

Die M-Zapfen und L-Zapfen reagieren auf sehr ähnliche Wellenlängen. Durch genetische Fehler kommt es bei den Betroffenen dazu, dass entweder M- oder L-Zapfen fehlen oder nicht richtig funktionieren. Das hat keine erheblichen Auswirkungen auf den wahrnehmbaren Farbbereich, allerdings kann es so dazu kommen, dass rot und grün zu derselben Farbe interpretiert werden und so nicht unterschieden werden können.

### **Aufgabe 10**

Die retinale Größe des Mondes können wir auf der gesamten Laufbahn in einer Nacht als konstant ansehen. Ebenso die Distanz von der Linse zur Retina im Auge. Da sich die wahrgenommene Größe trotzdem unterscheidet, muss es Unterschiede in der wahrgenommenen Distanz geben. Das kann man so erklären:

Befindet sich der Mond am Horizont, gibt es viele Anhaltspunkte in der Land-

schaft, die darauf hinweisen, dass der Mond sehr weit weg ist. Befindet sich der Mond am freien Nachthimmel, gibt es solche Anhaltspunkte nicht und die Distanz muss geraten werden. Diese intuitive Schätzung fällt meistens zu klein aus, sodass damit nach der Formel  $H \sim B \cdot D$  auch die wahrgenommene Größe des Mondes kleiner ist.

### Aufgabe 11

Die Schallgeschwindigkeit bei 20 °C ist  $v = 343 \text{ m/s}$ . Die Entfernung der beiden Personen beträgt  $s = 500 \text{ m}$ . Also berechnet sich die Zeit durch:

$$t = \frac{s}{v} = \frac{500 \text{ m}}{343 \text{ m/s}} \approx 1,458 \text{ s}$$

Dieselbe Formel kann auch für die Zeit des Lichtes verwendet werden. Dazu rechnen wir die Lichtgeschwindigkeit in  $\text{m s}^{-1}$  um:  $v_2 = 300\,000\,000 \text{ m/s}$ . Jetzt setzen wir ein:

$$t = \frac{s}{v_2} = \frac{500 \text{ m}}{300\,000\,000 \text{ m/s}} \approx 1,668 \cdot 10^{-6} \text{ s} = 1,668 \text{ } \mu\text{s}$$

Während das Bild des Händeklatschens Person B fast sofort erreicht, braucht es knapp 1,5 s, bis der dazugehörige Schall die 500 m überwunden hat.

### Aufgabe 12

Die Leistung berechnet sich folgendermaßen:

$$\text{Leistung} = \frac{\text{Zeit}_{\text{Kontrollgruppe}} - \text{Zeit}_{\text{Transfergruppe}}}{\text{Zeit}_{\text{Kontrollgruppe}}} \cdot 100$$

Also setzen wir ein:

$$\text{Leistung}_{\text{Methode 1}} = \frac{15 \text{ min} - 10 \text{ min}}{15 \text{ min}} \cdot 100 \approx 33,33 \%$$

$$\text{Leistung}_{\text{Methode 2}} = \frac{15 \text{ min} - 5 \text{ min}}{15 \text{ min}} \cdot 100 \approx 66,67 \%$$

$$\text{Leistung}_{\text{Kontrollgruppe}} = \frac{15 \text{ min} - 15 \text{ min}}{15 \text{ min}} \cdot 100 \approx 0 \%$$

Also hat Methode 2 die höchste Leistung.

### Aufgabe 13



1. Am Sonntag scheint die Sonne.  
Die richtige Deduktion ist „Es ist Sonntag. Also scheint die Sonne.“. Bei der anderen Aussage hätte man vorausgesetzt, dass die Sonne nur an Sonntagen scheint. Es könnte aber genauso gut ein anderer Tag sein, an dem die Sonne scheint.  
  
In der echten Welt gibt es natürlich auch Sonntage, an denen die Sonne nicht scheint. Die Deduktion ist aber grundsätzlich unabhängig von der inhaltlichen Korrektheit der Aussage. Es geht nur um formale Folgerbarkeit.
2. An manchen Montagen scheint die Sonne.  
Keine der beiden Aussagen ist eine gültige Deduktion. Bei „Es ist Montag. Also scheint die Sonne“ würden wir annehmen, dass an jedem Montag die Sonne scheint. Es könnte sich aber auch um einen Montag handeln, an dem gerade keine Sonne scheint. Bei der Aussage „Es scheint die Sonne. Also ist Montag“ machen wir den gleichen Fehler wie bei 1.
3. Alle geraden Zahlen sind durch 2 teilbar.  
Die einzige gültige Deduktion ist „Die Zahl 10 ist gerade. Also ist 10 durch 2 teilbar“. Durch unser Verständnis der Mathematik wissen wir natürlich, dass auch die anderen beiden Aussagen gelten. Sie sind trotzdem keine gültigen Deduktionen, da die Folgerung nicht aus der logischen Regel hervorgeht. Wenn wir nur wissen, dass alle geraden Zahlen durch 2 teilbar sind, könnte es immer noch sein, dass einige ungerade Zahlen ebenfalls durch 2 teilbar sind. Ebenso könnte es durch 2 teilbare Zahlen geben, die nicht gerade sind.  
  
Damit alle Aussagen korrekte Deduktionen wären, müsste man die logische Regel so formulieren: „Genau alle gerade Zahlen sind durch 2 teilbar“.

## Aufgabe 14

Lösung zur Menu-Auswahl

## Aufgabe 15

Im Human Action Cycle gibt es 7 Schritte. Diese sehen hier so aus:

1. Ziel: Das Ziel ist es, das Icon zu löschen.
2. Intention: Entscheiden zwischen Drag and Drop, Kontextmenu, Tastenkürzel, ... (im Folgenden betrachten wir Drag and Drop).
3. Handlungen: Maus finden, Maus ergreifen, Mauscursor finden, Mauscursor zum Icon bewegen, linke Maustaste drücken, Mauscursor zum Papierkorb bewegen, linken Mauscursor loslassen.

4. Ausführen der Handlungssequenz aus Schritt 3.
5. Wahrnehmen der Änderung: Man sieht während der Aktion eine Änderung. Danach sieht man, dass das Icon verschwunden ist.
6. Interpretation: Da das Icon verschwunden und der Papierkorb gefüllt ist, ist das Icon gelöscht.
7. Vergleich mit Ziel: Das Icon wurde erfolgreich gelöscht.

### **Aufgabe 16**

1. Gulf of Execution (Schritt 3). Wenn dem Benutzer nicht klar ist, in welcher Reihenfolge einzelne Aktionen ausgeführt werden müssen, liegt offenbar eine Unkenntnis bei der Bedienung des Systems vor. Dies entspricht Schritt 3 im HAC.
2. Gulf of Execution (Schritt 4). Zu kleine Icons fallen in die Kategorie von umständlicher Bedienung. Dieser Fehler tritt in Schritt 4 im HAC auf.
3. Gulf of Evaluation (Schritt 5, bzw. Schritt 6). Das Wahrnehmen des Ergebnisses findet in Schritt 5 statt. Damit scheitert hier ein nicht lesbares oder unerkennbares Ergebnis. Ein unverständliches Ergebnis kann zwar erkannt werden, kann dann aber nicht interpretiert werden und scheitert somit in Schritt 6.
4. Gulf of Evaluation (Schritt 7). Ist das Ergebnis unerwartet, kann dies erst durch einen Vergleich mit der Zielsetzung aus Schritt 1 festgestellt werden. Das passiert in Schritt 7.

### **Aufgabe 17**

1. Hier liegt ein Capture Error vor. Offenbar ist das Eintippen von „twitter“ so im Gedächtnis verankert, sodass man ohne Nachzudenken diese Zeichen eintippt, bevor man merkt, dass dies die falsche Seite ist.
2. Hier liegt ein Mode Error vor. Das Textfeld befindet sich durch die Feststelltaste im falschen Modus.
3. Hier liegt ein Mode Error vor. Die Kamera befindet sich anscheinend im Fotomodus und nicht im Videomodus.
4. Hier liegt ein Description Error vor. Das Feld „Benutzername“ ist offenbar das falsche Zielobjekt für die Eingabe des Passwortes.

5. Hier liegt ein Scheitern vor, das durch die mangelnde Fähigkeit oder eine zu hohe Schwierigkeit der Aufgabe entstanden ist.

### Aufgabe 18

### Aufgabe 19

### Aufgabe 20

- Ziele sollten nicht zu klein sein. Die Ziele müssen erkannt und gefunden werden.
  - Nach Fitts' Law sind kleine Ziele sehr schwer zu treffen. Daher wirken sich kleine Ziele nachteilig auf die Benutzerinteraktion aus.
  - Das Ergreifen und Finden von Zielen muss schnell erfolgen, da sich sonst im KLM-GOMS mehr Ms befinden. Dadurch stockt der Arbeitsablauf.
- Bei fortlaufenden Aktionen sollten Ziele nahe beieinander sein.
  - Nach einer Aktion befindet sich die Maus bei dem Auslöser der Aktion. Um nun die Zeit bis zur nächsten Aktion gering zu halten, sollte der nächste Auslöser möglichst nahe gelegen sein, sodass die Bewegungszeit nach Fitts' Law klein bleibt.
- Häufig gesuchte Ziele sollten möglichst immer an der gleichen Stelle sein.
  - Das Suchen und Finden von Zielen manifestiert sich durch Ms im KLM-GOMS. Befindet sich ein häufig gesuchtes Ziel immer an der gleichen Stelle, wird aus dem mentalen Akt eine Routinebewegung und die Ms verschwinden.
- Möglichst wenig weit entfernte Ziele.
  - Weit entfernte Ziele brauchen mehr Zeit, um getroffen zu werden (nach Fitts' Law). Zu viele weit entfernte Ziele sorgen für lange Wartezeiten zwischen zwei Aktionen und unnötig viel Mausbewegung.

### Aufgabe 22

Das Gehirn hat heute etwa 20 mal so viele Nervenzellen wie ein Computer Transistoren. In 1,5 Jahren halbiert sich dieser Faktor. In 3 Jahren ist der Faktor nur noch bei 5 usw. Dies kann durch die folgende Formel ausgedrückt und berechnet werden:

$$\log_2(20) \cdot 1,5 \approx 6,5$$

Also wird ein alltäglicher Computer in knapp 7 Jahren mehr Transistoren haben, als das menschliche Gehirn Nervenzellen hat.

### Aufgabe 23

#### Dies ist ein Link

Dies ist ein Link, der durch die blaue Farbe und die Unterstreichung offenkundig als solcher erkennbar ist. Damit ist dies eine erkennbare Affordanz.

#### **Dies ist ein Link**

Dies ist ein Link, der sich nicht vom normalen Text unterscheidet. Nach Gibson ist die Affordanz des Links zwar gegeben, nach Norman allerdings nicht. Damit handelt es sich um eine hidden affordance.

#### Dies ist kein Link

Normaler Text, der aussieht wie ein Link suggeriert dem Benutzer, dass er auf den Text wie auf einen Link klicken kann. Tatsächlich ist aber diese Affordanz nur nach Norman gegeben und nicht nach Gibson (da es ja kein echter Link ist). Also ist dies eine hidden affordance.

### Aufgabe 24

- Minimum: 0
- Maximum: 5
- Mittelwert:

$$\mu = \frac{0 + 1 + 1 + 2 + 3 + 4 + 4 + 4 + 5}{9} \approx 2.67$$

- Standardabweichung:

$$\sigma = \sqrt{\frac{(0 - 2.67)^2 + \dots + (5 - 2.67)^2}{9}} \approx 1.63$$

- Median: 3
- Modus: 4