

Informatik
Hauptcampus

H O C H
S C H U L E
T R I E R

Eine Einführung in mathematische Modelle der
biologischen Nervenzelle als Bausteine künstlicher
neuronaler Netze und deren Anwendung im
Gesundheitswesen

Thorsten Suckow-Homberg

Aachen, 28.09.2023

Eidesstattliche Erklärung

Hiermit versichere ich, die vorliegende Zulassungsarbeit selbstständig und nur unter Verwendung der von mir angegebenen Quellen und Hilfsmittel verfasst zu haben. Dies umfasst auch die Verwendung von KI-Anwendungen, die nach bestimmten Vorgaben automatisiert Inhalte erstellen können. Sowohl inhaltlich als auch wörtlich entnommene sowie von KI-Anwendungen generierte Inhalte wurden als solche kenntlich gemacht.

Die Arbeit wurde bisher keiner anderen Prüfungsbehörde vorgelegt und auch noch nicht veröffentlicht.

Thorsten Suckow-Homberg
Aachen, 28.09.2023

Kurzfassung

Neuronale Netze sind in den vergangenen Jahren ein wesentlicher Bestandteil von Forschung und Anwendung in der Medizin geworden.

Die vorliegende Arbeit stellt einige solcher Netze sowie deren theoretischen Unterbauten vor und trägt hierfür Ergebnisse aus Wissenschaft und Forschung zusammen.

Der erste Teil konzentriert sich zunächst auf die Fachlichkeit, die Grundlage künstlicher neuronaler Netze ist: Das biologische Neuron mit seinen komplexen biochemischen Vorgängen, die Signalverarbeitung und -weiterleitung in einem Netz aus Nervenzellen ermöglichen.

Darauf aufbauend bietet der zweite Teil eine Einführung in das mathematische Gerüst des McCulloch-Pitts-Neurons und Rosenblatt-Perzeptrons, zwei frühe Modelle künstlicher Neuronen, die wesentlich zur Entwicklung künstlicher neuronaler Netze beigetragen haben. Hierfür werden die Forschungsarbeiten der Autoren in Auszügen vorgestellt, und Anwendungsbeispiele beleuchten die unterschiedlichen Eigenschaften hinsichtlich Statik des McCulloch-Pitts-Neurons und Anpassungsfähigkeit des Rosenblatt-Perzeptrons.

Im dritten Teil werden einige wegweisende Architekturen und Algorithmen künstlicher neuronaler Netze vorgestellt, darunter Backpropagation sowie Faltungsoperationen, die in den tiefen Netzen Anwendung finden, die heutzutage u.a. in der Medizin große Erfolge vorweisen können.

In diesem Zusammenhang bezeugen im vierten und abschliessenden Teil ausgewählte Forschungsarbeiten und -Ergebnisse aus den Bereichen Gesundheitswirtschaft, der Pharmaforschung sowie der Diagnostik und Therapie die Leistungsfähigkeit künstlicher neuronaler Netze.

Inhaltsverzeichnis

1	Einführung	1
2	Signalweiterleitung im biologischen Neuron	2
2.1	Ionenkonzentrationen und Membranspannungen	3
2.2	Das Aktionspotenzial	6
2.2.1	Auslösung eines Aktionspotenzials	7
2.2.2	Signalweiterleitung über das Axon	9
2.3	Synaptische Übertragung und Integration	9
2.3.1	Synaptische Übertragung	9
2.3.2	Synaptische Integration	10
3	Künstliche Neuronen	12
3.1	Das McCulloch-Pitts-Neuron	12
3.1.1	Der Kalkül	12
3.1.2	Anwendung des Modells	13
3.1.3	Aktivierungs- und Eingabefunktion	14
3.1.4	McCulloch-Pitts-Netz als Graph	16
3.1.5	Implementierung von Booleschen Funktionen	16
3.2	Das Rosenblatt-Perzeptron	21
3.2.1	Das Perzeptron - ein linearer Klassifizierer	21
3.2.2	Das Modell	21
3.2.3	Lineare Trennbarkeit	23
3.2.4	Die Lernregel	25
3.2.5	Die XOR-Funktion	28
4	Künstliche Neuronale Netze	30
4.1	Renaissance der neuronalen Netze	30
4.2	Mehrschichtige neuronale Netze	31
4.2.1	Backpropagation	31
4.2.2	Hopfield-Netz	33
4.2.3	Neocognitron	35
4.2.4	Convolutional Neural Networks	36

5	Neuronale Netze im Gesundheitswesen	39
5.1	Automatisierung in der Gesundheitswirtschaft	40
5.2	Pharmaforschung	40
5.3	Diagnostik	41
5.4	Therapie und Prognose	43
6	Zusammenfassung und Bemerkung	45
	Anhang A Begriffe und Ergänzungen	51
A.1	Aktionspotenzial	51
A.1.1	Übertragungsgeschwindigkeit von Signalen in Nervenzellen	52
A.2	ATPasen	53
A.3	Dendriten	53
A.4	Diffusion	53
A.5	Effektoren	53
A.6	Expertensysteme	54
A.7	Goldman-Gleichung	54
A.8	KI Winter	55
A.9	LeNet	55
A.10	Lernrate	56
A.11	Neurotransmitter und ihre Rezeptoren	56
A.12	Perzeptron	57
A.13	Shunting Inhibition	58
A.14	Soma	58
A.15	Umkehrpotenzial	58
	Anhang B Personen	60
B.1	Bernstein, Julius	60
B.2	Hebb, Donald	61
B.3	Hodgkin, Alan und Huxley, Andrew	62
B.4	Hubel, David und Wiesel, Torsten	63
B.5	Loewi, Otto	63
B.6	McCulloch, Warren und Pitts, Walter	64
B.7	Nernst, Walther	65
B.8	Wiener, Norbert	66
	Anhang C Die Paradoxe Hitzeempfindung als MCP-Netz	67
	Anhang D Perzeptron-Implementierung in Python	72

Einführung

Der Einsatz künstlicher neuronaler Netze als Teilgebiet der Künstlichen Intelligenz konnte in der Medizin in den letzten Jahren beachtliche Erfolge vorweisen. Diese Arbeit stellt einige dieser Erfolge vor. Sie möchte darüber hinaus eine Einführung in Themen und Aspekte aus den Bereichen Neurobiologie, Mathematik und Informatik bieten, um ein erstes Verständnis für Modelle und Verfahren zu schaffen, die zu diesen Erfolgen beigetragen haben.

Wir beschäftigen uns deshalb zunächst mit dem Neuron als “strukturelle und funktionelle Einheit des Nervensystems“ [SD07, 42]. Im Hinblick auf neuronale Netze ist für uns das natürliche Vorbild der Signalweiterleitung in Nervenzellen von Interesse, weshalb wir uns eingehend mit der Zellmembran und dem *transmembranalen Transport* von Ionen beschäftigen: Es wird sich zeigen, dass Änderungen der Nervenzellmembran-Eigenschaften hinsichtlich ihres Potentials grundlegend sind für die Ausschüttung von hemmenden oder erregenden Neurotransmittern, die von den Rezeptoren nachgelagerter Nervenzellen weiterverarbeitet und dort gemäß eines Alles-oder-Nichts-Prinzips erneut Signale auslösen können.

Dies bildet die Grundlage für die Neuronenmodelle von McCulloch und Pitts sowie Rosenblatt im zweiten Teil der Arbeit, in dem der Schwerpunkt vor allem auf der mathematischen Modellierung von Eingabe- und Ausgabefunktion sowie deren Anwendung liegt: Ein nicht unwesentlicher Teil der zu erarbeitenden Fachlichkeit beschränkt sich also nicht bloß auf die hochkomplexen biochemischen Vorgänge, die in der biologischen Nervenzelle bei Signalverarbeitung und -weiterleitung stattfinden. Einige diese Vorgänge werden in diesem Abschnitt in Formeln und Gleichungen der Aussagenlogik bzw. analytischen Geometrie überführt, damit sie in der Informatik Anwendung finden können.

Im dritten und vierten Teil werden darauf aufbauend Architektur und Algorithmen einiger künstlicher neuronaler Netze sowie deren überaus erfolgreiche Anwendung im Gesundheitswesen vorgestellt.

Signalweiterleitung im biologischen Neuron

Eingehende Signale empfängt eine Nervenzelle über seine **Dendriten**: Baumförmige Fortsätze, die um den Zellkörper - das **Soma** - herum gelagert sind¹. Sie fungieren *postsynaptisch* und empfangen über **Rezeptoren** afferente² Signale in Form von **Neurotransmittern**³ (siehe Abbildung 2.1).

Dendriten leiten die Signale weiter an das Soma, in dem sich, durch die **Neuronenmembran**⁴ von der Umgebung getrennt, **Zytosol** befindet: Eine salzige, wässrige Flüssigkeit mit einem hohen Anteil von Kalium⁵.

Ob das Neuron Signale weiterleitet, entscheidet sich in der Nähe des **Axonhügels**⁶: Hier entspringt das **Axon**⁷, das in einer "salzigen extrazellulären Flüssigkeit mit hoher Leitfähigkeit" liegt [BCP18, 61]⁸. Reicht die Integration der an den postsynaptischen Endigungen eingehenden Signale aus, kann eine Depolarisation⁹ der Membran an dieser Stelle¹⁰ erfolgen. Wird ein gewisser **Schwellenwert** übertroffen, wird ein **Aktionspotenzial** ausgelöst (vgl. [BCP18, 142 f.]), was in den präsynaptischen Endigungen die **Exozytose**¹¹ auslöst.

¹ siehe Anhang A.14

² "afferre" (lat.): herbeibringen, melden, bringen

³ vgl. [Eil19, 61]. Oft stehen tausende Neuronen in Verbindung mit den Dendriten eines einzelnen Neurons (vgl. [SD07, 42]). Die Anzahl der Neuronen im menschlichen Gehirn beziffert sich auf mindestens 10^{11} (vgl. [Kan+13, 175]).

⁴ Membrandicke ca. 5 nm (vgl. [FE19, 66])

⁵ vgl. [BCP18, 29]. Siehe hierzu auch "Ionenkonzentrationen" in Tabelle 2.1.

⁶ ca. 20 - 50 μm vom Soma entfernt (vgl. [Jon19, 77])

⁷ "axon" (lat.): Achse; Axone können sich im menschlichen Körper über Entfernungen von bis zu über 1m ausstrecken (vgl. [BCP18, 28])

⁸ *Bear et al.* führen das Axon metaphorisch mit einer Telefonleitung zusammen (siehe [BCP18, 43]). Aufgrund der signalempfangenden Eigenschaften und der dünnen Spitzen der Dendriten werden diese auch mit Antennen verglichen (siehe [BCP18, 28])

⁹ *Depolarisation* bezeichnet die Verringerung des Membranpotenzials von einem negativen Wert auf einen weniger negativen oder gar einen positiven Wert (vgl. [McC13, 95 f.]

¹⁰ vgl. [Eil19, 61]

¹¹ bezeichnet den Stofftransport aus der Zelle heraus.

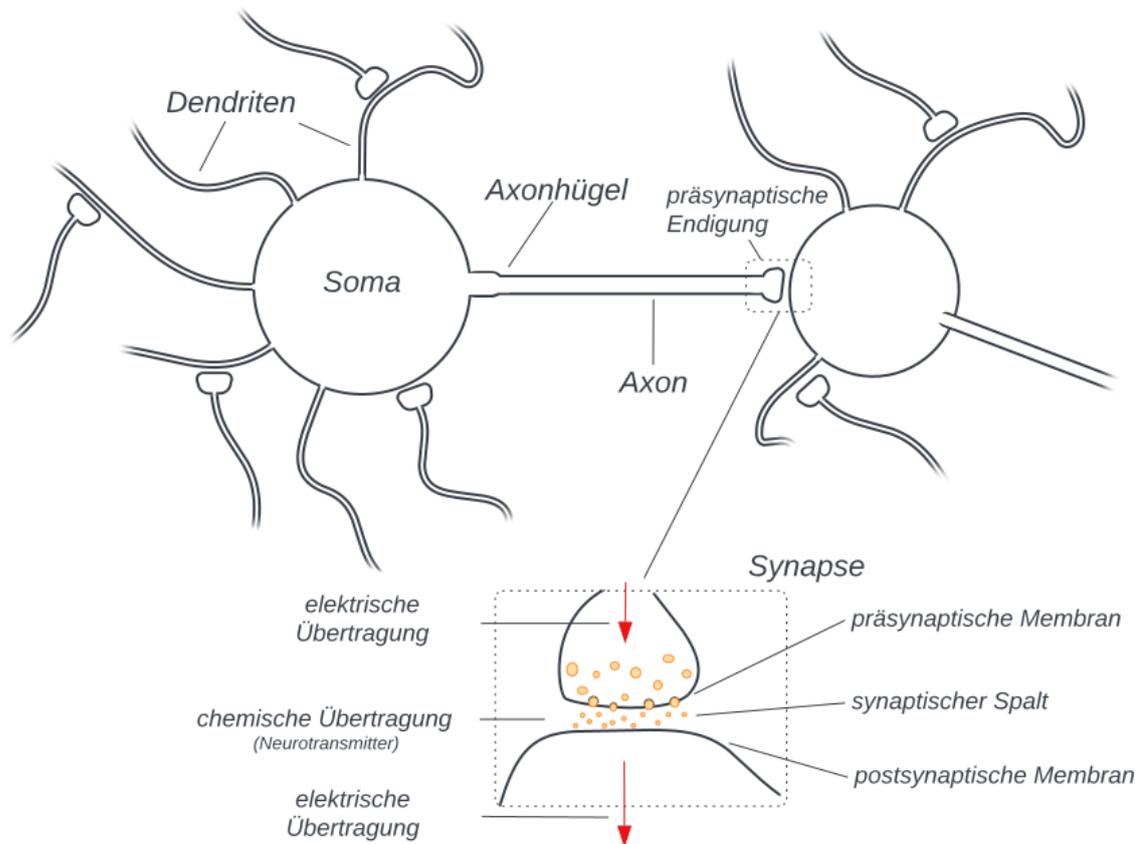


Abb. 2.1. Schematische Darstellung einer Nervenzelle. (Quelle: in Anlehnung an [SD07, 43, Tafel 2.1])

Die Dendriten leiten afferente Signale zum Soma, dem Zellkörper. Das Axon leitet ein efferentes („efferre“ (lat.): hinaustragen, mitnehmen) Nervensignal über präsynaptische Endigungen (auch „Axonterminale“) an (häufig weit entfernte) Effektoren wie Muskeln, Drüsen oder nachgeschaltete Neuronen weiter.

2.1 Ionenkonzentrationen und Membranspannungen

Ein Neuron weist *in Ruhe*¹² eine ungleiche Ionenverteilung zwischen der durch die Zellmembran getrennten **intrazellulären Flüssigkeit** (das Zytosol, im folgenden auch IZF) und der **extrazellulären Flüssigkeit** (im folgenden EZF) auf. In der IZF befinden sich mehr positiv geladene Natrium-Ionen (Na^+), und in der EZF mehr positiv geladene Kalium- und Calcium-Ionen (K^+ und Ca^{2+}) sowie mehr negativ geladene Chlorid-Ionen (Cl^-).

Das **Membranpotenzial** des Neurons wird durch die Verteilung der Ionen in der IZF und EZF bestimmt: In der Membran befinden sich **Ionenkanäle**, von de-

¹² vgl. **Membranpotenzial**: „die Spannung an der Nervenzellmembran zu einem beliebigen Zeitpunkt“ [BCP18, 70]; **Ruhepotenzial**: „the electrical potential across the membrane in the absence of signaling“ [Kan+13, 126]

Ion	Konzentration EZF (mmol/l)	Konzentration IZF (mmol/l)	Verhältnis
K^+	5	100	1 : 20
Na^+	150	15	10:1
Ca^{2+}	2	0,0002	10000 : 1
Cl^-	150	13	11,5 : 1

Tabelle 2.1. Ionenkonzentration eines Neurons in Ruhe (Quelle: [BCP18, 75, Abb. 3.15])

nen viele **selektiv permeabel**¹³ sind. Daneben existieren **Ionenpumpen** wie die Ca^{2+} - und Na^+ - K^+ -ATPasen. Sie sorgen dafür, dass im Neuron laufend Ca^{2+} und Na^+ aus der Zelle und K^+ in die Zelle befördert wird (vgl. [SD07, 44]). Zusammen mit den selektiv permeablen Ionenkanälen entstehen so die Ionenkonzentrationen in Tabelle 2.1.

Wenn kein **postsynaptisches Potenzial** (PSP) wirkt und das Neuron selber keinen Impuls abgibt, liegt das Ruhepotenzial V_r der Zelle zwischen $-70mV$ und $-90mV$ ¹⁴: Das Zytosol weist entlang der Membranoberfläche in der IZF eine negative Ladung auf¹⁵. Diese **Membranspannung** V_m wird durch eine ungleiche Ionenverteilung bewirkt, verursacht durch die Ladung der Teilchen in der IZF und EZF in Membrannähe¹⁶ (siehe Abbildung 2.2).

In Ruhe ist die Leitfähigkeit der Membran für Na^+ gering, für K^+ hingegen hoch (vgl. [SD07, 44]). K^+ -Ionen folgen ihrem **Konzentrationsgradienten** und gelangen über die Ionenkanäle in die EZF, bis die **Potenzialdifferenz** entlang der Neuronenmembran ausströmende K^+ -Ionen zurückhält: Wenn diese Differenz den Konzentrationsgradienten für K^+ kompensiert, ist das **Gleichgewichtspotenzial** erreicht, und es findet keine Nettoionenbewegung statt¹⁷.

¹³ von „*permeare*“ (lat.) durchwandern. Ein selektiv permeabler Kanal ist nur für bestimmte Ionen durchlässig (*ionenselektiv*). Kaliumkanäle sind durchlässig für K^+ -Ionen, Natriumkanäle durchlässig für Na^+ -Ionen usw. (vgl. [BCP18, 66]). Die Kanäle können über Änderungen in der Umgebung des Neurons geöffnet oder geschlossen werden, was auch als **Gating** bezeichnet wird (vgl. [Kan+13, 108])

¹⁴ vgl. [SD07, 47, Tafel 2.3 (A.1)]. *Bear et al.* geben $-65mV$ an ($1mV = 0,001V$), es wird von einer 40-mal höheren Ionenpermeabilität für K^+ gegenüber Na^+ ausgegangen (vgl. [BCP18, 74, Exkurs 3.2] sowie [BCP18, 70])

¹⁵ vgl. [BCP18, 61]. Das Membranpotenzial V_m ergibt sich als die Differenz der Spannungen V_{IZF} und V_{EZF} , wobei V_{IZF} die Spannung in der IZF und V_{EZF} die Spannung in der EZF ist. V_r ist dann gleich zu V_{IZF} : Wenn die Zelle in Ruhe ist, wird die Spannung in der EZF mit 0 festgelegt (vgl. [Kan+13, 127]).

¹⁶ vgl. [FE19, 66]. *Bear et al.* stellen fest: „Die negativen Ladungen im Inneren des Neurons und die positiven Ladungen außerhalb des Neurons ziehen sich in Richtung Zellmembran gegenseitig an, [...] Dementsprechend ist die negative Nettoladung im Inneren der Zelle nicht gleichmäßig im Cytosol verteilt, sondern an der Innenseite der Membran lokalisiert.“ [BCP18, 72, Punkt 2]

¹⁷ vgl. [BCP18, 72]. *Kandel et al.* schreiben hierzu: „the equilibrium potential of any ion that is present on both sides of a membrane permeable to that ion“ [Kan+13, 130]

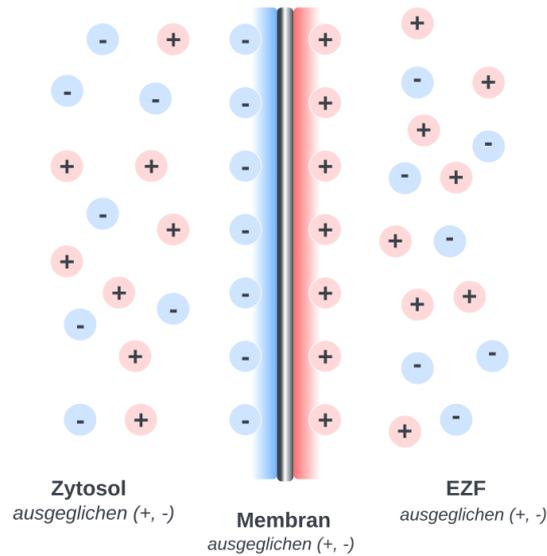


Abb. 2.2. Ionenverteilung im Zytosol und der EZF. (Quelle: in Anlehnung an [BCP18, 72, Abb. 3.13])

Aufgrund der elektrostatischen Anziehungskraft ziehen sich Anionen (neg. geladen) und Kationen (pos. geladen) in der Nähe der Membran gegenseitig an, es kommt zu einer negativen Spannung in Membrannähe (zwischen $-70mV$ und $-90mV$ in Ruhe). Dadurch ist die negative Ladung nicht gleichmässig im Zytosol verteilt. *Bear et al.* weisen darauf hin, dass trotzdem der größte Teil des Zytosols und der EZF elektrisch neutral ist (vgl. [BCP18, 72, Abb. 3.13])

Es gilt, dass sich das Membranpotenzial dem Gleichgewichtspotenzial desjenigen Ions annähert, für das die Membran besonders permeabel ist (vgl. [Kan+13, 145 f.]): Bestimmen lässt sich das Gleichgewichtspotenzial für individuelle Ionen mit der Nernst-Gleichung¹⁸ (s. Gleichung 2.1 sowie Tabelle 2.2):

Bear et al. definieren¹⁹:

$$E_{Ion} = 2,303 \times \frac{RT}{zF} \times \log_{10} \times \frac{[Ion]_{EZF}}{[Ion]_{IZF}} \quad (2.1)$$

Für eine Körpertemperatur von 37° lässt sich die Nernst-Gleichung für das Gleichgewichtspotenzial E_K wie folgt vereinfachen:

$$E_K = 61,54mV \times \log_{10} \frac{[K^+]_{EZF}}{[K^+]_{IZF}} \quad (2.2)$$

Mit den Werten aus Tabelle 2.1 ergibt sich somit

¹⁸ vgl. [FE19, 67].

¹⁹ siehe [BCP18, 74, Exkurs 3.2]

Variable / Konstante	Bedeutung
E_{ion}	Gleichgewichtspotenzial für das jeweilige Ion
R	Gaskonstante
T	absolute Temperatur
z	Ladungszahl des Ions
F	Faraday-Konstante
$[Ion]_{EZF}$	Ionenkonzentration ausserhalb der Zelle
$[Ion]_{IZF}$	Ionenkonzentration innerhalb der Zelle

Tabelle 2.2. Nomenklatur Nernst-Gleichung

E : *Equilibrium* ("Gleichgewicht"); *Faraday-Konstante*: elektrische Ladung eines Mols einfach geladener Ionen; 1 Mol = $6,02214076 \times 10^{23}$ Teilchen

$$E_K = 61,54mV \times \log_{10} \frac{1}{20} \approx -80mV \quad (2.3)$$

Zur Bestimmung von V_m einer für mehrere Ionen permeablen Membran kann die **Goldman-Gleichung**²⁰ verwendet werden:

$$V_r = \frac{RT}{F} \times \ln \frac{P_{Na} \times [Na^+]_{EZF} + P_K \times [K^+]_{EZF} + P_{Cl} \times [Cl^-]_{IZF}}{P_{Na} \times [Na^+]_{IZF} + P_K \times [K^+]_{IZF} + P_{Cl} \times [Cl^-]_{EZF}} \quad (2.4)$$

Eine weitere wichtige Eigenschaft der Membran ist die *elektrische Leitfähigkeit* g_{Ion} ; für sie gilt, dass sie proportional zu der Anzahl der offenen Ionenkanäle für *Ion* ist (vgl. [BCP18, 93]).

2.2 Das Aktionspotenzial

In Ruhe ist die Ladung in Membrannähe in der IZF negativ, in der EZF positiv. Durch eine schnelle Umkehrung dieser Verhältnisse ist eine Nervenzelle dazu in der Lage, ein Signal auszulösen (vgl. [BCP18, 86]). Hierzu bedarf es eines **Aktionspotenzials**, das durch Ionenbewegungen entsteht. Die Ionenbewegungen selber sind auf die Änderung des Membranpotenzials zurückzuführen, wodurch ein Öffnen oder Schliessen von ionenspezifischen Kanälen verursacht wird (vgl. [BCP18, 96]).

Kandel et al. führen 4 wichtige Eigenschaften des Aktionspotenzials auf (vgl. [Kan+13, 148 f.]):

²⁰ siehe Anhang A.7

1. Es gibt einen **Schwellenwert** für die Auslösung des Potenzials.
2. Das Aktionspotenzial ist ein **Alles-oder-Nichts** Ereignis.
3. Das Aktionspotenzial wird ohne Verlust weitergeleitet.
4. Nach dem Auslösen des Aktionspotenzial kommt es zu einer **Refraktärzeit**, in der zunächst kein weiteres Aktionspotenzial ausgelöst werden kann (**absolute Refraktärzeit**). Bevor V_r wieder erreicht ist, wird für eine kurze Dauer ein stärkeres Signal benötigt, um das Aktionspotenzial erneut auszulösen (**relative Refraktärzeit**)²¹.

2.2.1 Auslösung eines Aktionspotenzials

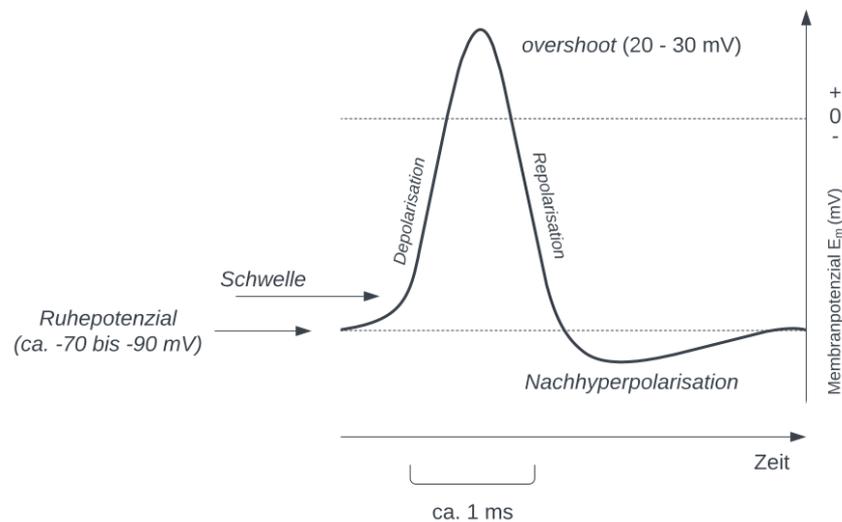


Abb. 2.3. Aktionspotenzial. (Quelle: in Anlehnung an [SD07, 47, Tafel 2.3])

Die **Initiationszone** des Aktionspotenzials ist der Axonhügel. Hier findet sich eine besonders dichte Anhäufung von spannungsabhängigen Na^+ -Kanälen²², deren Aktivierungskurve um ca. $10mV$ zu den negativen Membranpotenzialen verschoben ist, was die Initiierung des Signals begünstigt (vgl. [Jon19, 77]).

Damit ein Aktionspotenzial ausgelöst werden kann, muss die Membran nahe der Initiationszone über ihren **Schwellenwert**²³ V_t ($> V_r$) depolarisiert werden (vgl. [BCP18, 111]).

²¹ beide Zeitabschnitte sind Teil der **Nachhyperpolarisation**, dargestellt in in Abbildung 2.3. Bei der "Hyperpolarisation" ändert sich das Membranpotenzial auf negative Werte in Richtung V_r und darunter und wird in diesem Kontext als Hemmung verstanden.

²² pro Quadratmikrometer (μm^2) kann eine Membran viele tausend Natriumkanäle enthalten (vgl. [BCP18, 99])

²³ "das kritische Niveau der Depolarisation, das überschritten werden muss, um ein Aktionspotenzial auszulösen" [BCP18, 88] oder auch "Erregungsschwelle" [FE19, 69]

Als Schwellenwert wird das Membranpotenzial bezeichnet, bei dem die Permeabilität für Na^+ größer als für K^+ ist (vgl. [BCP18, 103])²⁴; V_t liegt bei ca. $-50mV$, unabhängig vom Typ des Neurons (vgl. [Jon19, 75]). Damit sich V_r an V_t nähert, bedarf es einer Erregung der Zelle durch postsynaptische Potenziale (vgl. [FE19, 69]) oder “eine aus der Umgebung weitergeleitete (elektrotonische) Erregung“ [SD07, 46].

Die Stärke der Erregung der Zelle ist entscheidend für das Auslösen eines Aktionspotenzials. Es muss mehr Na^+ in die Zelle einströmen, als K^+ aus der Zelle ausströmen kann (vgl. [FE19, 69]). Versuche zeigen, dass die Potenzialänderung der Membran in einem Bereich von $-80mV$ zu $-65mV$ in dieser Hinsicht kaum Änderung bewirkt (vgl. [BCP18, 99]). g_{Na} erhöht sich, aber wenn V_r nicht erreicht wird, bleibt es bei dieser “lokalen Antwort“ [SD07, 46]. Erst eine Depolarisation der Membran hin über diesen Wert²⁵ kann die **Initiationsphase**²⁶ des Aktionspotenzials einleiten: Es öffnen sich mehr Natrium-Kanäle, und durch die negative Ladung der Membran-Innenseite gibt es eine starke elektrochemische Triebkraft für Na^+ -Ionen (vgl. [BCP18, 103]). Die Triebkraft ist zu diesem Zeitpunkt die Differenz des Membranpotenzials V_m und E_{Na} . Bei $-60mV$ beträgt die Triebkraft²⁷:

$$V_m - E_{Na} = -60mV - 61,54mV = -121,54mV \quad (2.5)$$

Da sich durch zunehmende Öffnung der Natrium-Kanäle g_{Na} erhöht, strömen aufgrund der hohen Triebkraft für Na^+ mehr Na^+ -Ionen in das Zellinnere. Es kommt zu einem **Rückkopplungseffekt**, denn die zunehmend weniger negative Membranspannung öffnet weitere Natrium-Kanäle, die Leitfähigkeit der Membran wird weiter erhöht und es kommt zu einem exponentiellen Anstieg der Na^+ -Konzentration in der IZF. Der “explosionsartige“ Natriumeinstrom bewirkt die Depolarisation der Membranspannung und das Aktionspotenzial wird ausgelöst (vgl. [FE19, 69])²⁸.

Die Stärke des Signals selber ist unabhängig von dem Wert, der zu der überschweligen Reizung des Neurons geführt hat: Amplitude und Zeitverlauf des Signals im Axon hängen nicht von Intensität und Dauer des Reizes ab (vgl. [Jon19, 75]). Entweder kommt es zu einem Aktionspotenzial, oder es bleibt bei einer lokalen Antwort. Aus diesem Grund funktionieren Aktionspotenziale nach dem **Alles-oder-Nichts-Prinzip**²⁹:

“A fraction of a millivolt can be the difference between a subthreshold stimulus and a stimulus that generates a full-sized action potential.“ ([Kan+13, 157])

²⁴ V_r ist in Ruhe nahe an E_K .

²⁵ ab $-60mV$ gehen die Natrium-Kanäle in den Offen-Zustand über (vgl. [FE19, 69]).

²⁶ siehe [FE19, 68]

²⁷ vgl. [Fak19, 39]

²⁸ siehe hierzu auch Anhang A.1

²⁹ vgl. [BCP18, 89]. Außerdem: “Im Computerzeitalter bezeichnet man das axonale Aktionspotenzial auch als ‘digitales’ Signal.“ [Jon19, 75]. Frank bietet eine Übersicht über die Erforschung des Alles-oder-Nichts-Prinzip in [Fra94]. Dort wird *Lucas* erwähnt, der bereits 1905 in [Luc05] die Kontraktionen von Muskelfasern unter diesem Gesichtspunkt untersucht hat (vgl. [Fra94, 210])

2.2.2 Signalweiterleitung über das Axon

Kurz nach der Aktionspotenzialbildung befinden sich daran beteiligte Membrane in der Refraktärphase, ihre Na^+ -Kanäle sind inaktiviert. Somit pflanzt sich das Aktionspotenzial nur in eine Richtung fort³⁰ (vgl. [BCP18, 106]). Die Fortleitung geschieht in Richtung der Axonterminal, wo sich die präsynaptischen Endigungen befinden und das elektrische Signal in chemische Signale umgewandelt werden (siehe Abbildung 2.1).

2.3 Synaptische Übertragung und Integration

2.3.1 Synaptische Übertragung

Neben elektrischen Synapsen, bei denen der Signalaustausch durch einen direkten Stromfluss über sogenannte *gap junctions* und deren ionenleitfähige Verbindungen³¹ passiert (vgl. [BCP18, 119]), erfolgt die Signalweiterleitung im Gehirn überwiegend auf chemische Weise (vgl. [BCP18, 121 ff.]).

Chemische Synapsen sind nicht direkt miteinander verbunden. Zwischen ihnen existiert ein Spalt, der ca. 20 - 40nm breit ist (vgl. [Kan+13, 184]), und in der sich eine Matrix aus extrazellulären Proteinen befindet³², die den Synapsenspalt überbrückt (vgl. [BCP18, 122]). Die Übertragung von Signalen erfolgt über Exozytose: Botenstoffe (*Neurotransmittern*) diffundieren aus den präsynaptischen Endigungen in diesen Spalt (vgl. [BCP18, 122]). Rezeptoren an den postsynaptischen Endigungen wandeln die Botenstoffe in hemmende oder erregende Signale um, die dann von der postsynaptischen Zelle nach dem Alles-oder-Nichts-Prinzip integriert werden³³.

Die durch das Aktionspotenzial ausgelöste Depolarisation der Membran an den Axonterminalen bewirkt eine Öffnung spannungsgeladener Calcium-Kanäle (vgl. [Kan+13, 184]). Durch die ungleiche Ca^{2+} -Ionenkonzentration zwischen der EZF und IZF (10.000 : 1, siehe Tabelle 2.1) entsteht eine hohe Triebkraft für Ca^{2+} : Nach Gleichung 2.1 ergibt sich für das Gleichgewichtspotenzial für Ca^{2+}

$$E_{Ca^{2+}} = 123,08mV \quad (2.6)$$

und bei einem Membranpotenzial von $V_m \sim 20mV$ durch Depolarisation liegt die Triebkraft für Ca^{2+} bei $\sim -100mV$:

$$V_m - E_{Ca^{2+}} = 20mV - 123,08mV = -103,08mV \quad (2.7)$$

³⁰ *orthodrome* Fortleitung. *Bear et al.* verweisen auf *antidrome* Fortleitung (entgegengesetzte Verlaufsrichtung), die in Experimenten ausgelöst werden kann [BCP18, 106]

³¹ "Konnexone"; vgl. [SD07, 50]

³² die *extrazelluläre Matrix*. Hierbei handelt sich um den Gewebeanteil im *Interzellularraum*, also der Raum ausserhalb der Zellen: "The extracellular matrix [...] surrounds all connective tissue cells providing mechanical support and physical strength to tissues, organs and the organism as a whole" [Aya+98, 3]

³³ siehe hierzu auch Anhang A.11

Die Calcium-Ionen strömen in das Innere der Zelle und lösen die Exozytose von **synaptischen Vesikeln**³⁴ aus, kleine, mit einer Membran von der IZF getrennte Strukturen von etwa 50nm Durchmesser, die mit Neurotransmittern gefüllt sind (vgl. [BCP18, 1000]). Im *synaptischen Endknöpfchen* befinden sich 100–200 von diesen Bläschen, die jeweils tausende Moleküle eines Neurotransmitters beinhalten (vgl. [Kan+13, 184]).

Der Calciumeinstrom in die präsynaptische Endigung löst die Verschmelzung dieser Bläschen mit der Zellmembran des Endknöpfchens an der sogenannten **aktiven Zone**³⁵ aus: Ein spezialisierter Abschnitt der Membran, der direkt gegenüber der **postsynaptischen Dichte** liegt^{36,37}

2.3.2 Synaptische Integration

Die Neurotransmitter eines einzelnen Vesikels lösen einen minimalen exzitatorischen oder inhibitorischen postsynaptischen Strom aus (**mEPSC** bzw. **mIPSC**, *miniature excitatory / inhibitory postsynaptic current*)³⁸.

Wesentlich für die Entstehung eines neuen Aktionspotenzials in der postsynaptischen Zelle ist die Verrechnung dieser Signale, die räumlich oder zeitlich eintreffen. Hierbei ist die **räumliche Summation** die Integration von vielen fast gleichzeitig eintreffenden Signalen mehrerer präsynaptischer Zellen, die sich in der Folgezelle zu einem **exzitatorischen postsynaptischen Potenzial** (EPSP) aufaddieren (vgl. [HS19a, 101]). Unter der **zeitlichen Summation** versteht man die in gewissen Abständen von ein und derselben Synapse hintereinander eintreffenden EPSPs, die jeweils das Membranpotenzial für nachfolgende Signale zum Schwellenwert hin verschieben³⁹.

Ob der Schwellenwert der postsynaptischen Zelle überschritten werden kann ist auch abhängig von dem Einfluss der **inhibitorischen Synapsen** auf die Zelle: Inhibitorische “Eingaben“, die hyperpolarisierend wirken, werden von den exzitatorischen Eingaben subtrahiert^{40,41} (vgl. [Kan+13, 225]).

³⁴ “*vesicula*“ (lat.): “Bläschen“

³⁵ *Bear et al.* beschreiben das Aussehen der aktiven Zone als “ein Feld winziger Pyramiden“ [BCP18, 123]; *Rohkamm und Kerner* erklären die direkte Einwirkung (ionotrop) und indirekte Einwirkung (metabotrop) der Neurotransmitter (siehe [RK18, 134])

³⁶ damit ist der Abschnitt der postsynaptischen Endigung gemeint, in der sich die Rezeptoren befinden (vgl. [HS19a, 96])

³⁷ Kurz nach Beginn der Transmitterübertragung über den synaptischen Spalt findet die **Endozytose** statt, ein Recyclingprozess, in dem die individuellen Vesikelmembranen wiederhergestellt und mit Neurotransmitter erneut aufgefüllt werden (vgl. [BCP18, 133]).

³⁸ EPSCs und IPSCs setzen sich aus einzelnen dieser mEPSC bzw. mIPSC zusammen und bilden die kleinste Einheit der postsynaptischen Ströme, weshalb man sie als *Quanten* bezeichnet [HS19a, 98]. Da EPSPs ein Vielfaches des Quantums sind, “das die Menge an Transmitter in einem einzigen Vesikel und die Anzahl der postsynaptischen Rezeptoren an der Synapse widerspiegelt“, nennt man sie “gequantelt“ [BCP18, 142]. Mit Hilfe der **Quantelungsanalyse** lässt sich die Anzahl der an einer synaptischen Übertragung beteiligten Vesikel bestimmen (vgl. [BCP18, 142]).

³⁹ vgl. [BCP18, 142] sowie [HS19a, 101]

⁴⁰ *Bear et al.* stellen in [BCP18, 146, Exkurs 5.6] die Rolle inhibitorischer Synapsen anschaulich dar.

⁴¹ Daneben ist auch der Abstand der Synapsen von der Initiationszone sowie die Eigenschaften der dendritischen Membran zu berücksichtigen. So kann die EPSP-Amplitude kleiner werden, wenn Strom

Das Zusammenspiel zwischen exzitatorischen und inhibitorischen Synapsen wird auch durch das **Umkehrpotenzial** V_{rev} bestimmt⁴²: Im Allgemeinen gilt, dass an erregenden Synapsen **unspezifische Kationenkanäle** öffnen, deren Umkehrpotenzial im Bereich von $0mV$ liegt: Hier kommt es zu einer Depolarisation. An hemmenden Synapsen öffnen Cl^- - oder K^+ -Kanäle und es gilt dort $V_{rev} \leq V_r$, wonach es meist zu einer Hyperpolarisation kommt (vgl. [HS19a, 100]).

Die Verrechnung von EPSP und IPSP (inhibitorisches postsynaptisches Potenzial) erfolgt nicht ausschließlich linear durch Summation - inhibitorische Synapsen können auch für einen "Kurzschluss" sorgen und somit ein EPSP um ein Vielfaches verkleinern (vgl. [Sil10, 477]). Man spricht dann von einer *Kurzschlusshemmung* (**shunting inhibition**)⁴³.

auf dem Weg zu dem Axonhügel durch die Membran verloren geht (vgl. [BCP18, 142 f.]). *Bear et al.* vergleichen die Dendriten diesbzgl. mit einem "löchrigen Gartenschlauch" (vgl. [BCP18, 143]).

⁴² siehe Anhang A.15

⁴³ siehe Anhang A.13

Künstliche Neuronen

3.1 Das McCulloch-Pitts-Neuron

Im Jahr 1943 veröffentlichen *Warren McCulloch* und *Walter Pitts*¹ ihre Arbeit [MP43]. Sie stellen darin ein auf **Aussagenlogik**² basierendes mathematisches Modell zur Erklärung von Signalverarbeitung und -weiterleitung im Nervensystem vor, und liefern damit grundlegende Ideen für die Kybernetik³ und die Von-Neumann-Rechnerarchitektur⁴. Ihre Arbeit leistet insgesamt einen wichtigen Beitrag für die Entwicklung *Künstlicher Intelligenz* (vgl. [Arb00, 1]).

3.1.1 Der Kalkül

In der **Principia Mathematica**⁵ stellen *Russel* und *Bertrand* die Mathematik auf ein Gerüst logischer Grundprinzipien (vgl. [She26, 225]). Inspiriert dadurch beschäftigt sich McCulloch mit der Frage, ob es nicht auch möglich sei, die Vorgänge im Gehirn durch Anwendung logischer Prinzipien erklärbar zu machen (vgl. [Arb00, 4]). Die Zweiwertigkeit des Alles-oder-Nichts-Prinzips (siehe Abschnitt 2.2) führt zu dem Kalkül:

The ‘all-or-none law’ of nervous activity is sufficient to insure that the activity of any neuron may be represented as a proposition. ([MP43, 100])

McCulloch und Pitts gehen für ihren Kalkül von folgenden physischen Eigenschaften einer Nervenzelle aus⁶ (vgl. [MP43, 101]):

¹ siehe Anhang B.6

² Die (zweiwertige) Aussagenlogik beschäftigt sich mit der Verknüpfung von Aussagen und deren Wertigkeit unter der Berücksichtigung der Wahrheitswerte “wahr“ und “falsch“ (vgl. [Rau08, 2]).

³ siehe Anhang B.8

⁴ Die Architektur zeichnet sich durch eine Zentraleinheit, der Speichereinheit und Ein-/Ausgabeeinheit aus, die über Busse verbunden sind (vgl. [OV00, 230])

⁵ Die **Principia Mathematica** [WR27] (PM) der Philosophen und Mathematiker Bertrand Russell (1872 - 1970) und Alfred Whitehead (1861 - 1947) ist ein Werk über die Grundlagen der Mathematik und erschien in drei Bänden zwischen 1910 und 1913.

⁶ Für ihren Kalkül entwerfen sie zwei Netze: Ein zyklensfreies Netz (vgl. [MP43, 101]) sowie ein Netz, in dem Zyklen vorhanden sind (vgl. [MP43, 108]), wobei die Frage nach Ursprung und Bedeutung von “loops“ - also Zyklen - in einem Netz von Nervenzellen McCulloch bereits seit 1928 beschäftigt: “The

1. Ein Neuron arbeitet auf Basis des Alles-oder-Nichts-Prinzip.
2. Eine gewisse Anzahl von Synapsen muss innerhalb einer bestimmten Zeit angeregt werden, um ein Neuron zu erregen.
3. Die einzige signifikante Verzögerung bei der Signalübertragung innerhalb des Nervensystems ist bedingt durch synaptische Verzögerungen.
4. Eine einzige hemmende Synapse verhindert die Erregung des betreffenden Neurons
5. Die Struktur des neuronalen Netzes ändert sich nicht.

3.1.2 Anwendung des Modells

Im folgenden wollen wir das Modell von McCulloch und Pitts anwenden. Hierzu überführen wir die für die Original-Arbeit aus der PM entnommene Notation für Aussagenlogik⁷ in eine geläufigere Form: Wir nutzen “ \wedge ” für *Konjunktion*, “ \vee ” für *Disjunktion* sowie “ \neg ” für die *Negation*. “ \equiv ” verwenden wir im Sinne der Äquivalenz. Für “McCulloch-Pitts-Neuron” verwenden wir im Folgenden abkürzend **MCP-Zelle**; für ein Netz, das aus MCP-Zellen besteht, einfach **MCP-Netz**.

In Anlehnung an das Original legen wir folgende Anforderungen für die Anwendung fest⁸

1. Das Verhalten des künstlichen Neurons ist zweiwertig, also *binär*. Das Neuron feuert, oder es feuert nicht. Im Sinne der Beziehung zwischen Eingabe und Ausgabe verstehen wir unter “feuern” den Wert 1. Wenn ein Neuron nicht feuert, wird dies durch den Wert 0 repräsentiert, der ebenfalls als Ausgabe vorliegt. In diesem Sinne produziert unser Modell also immer ein “Signal”⁹.
2. An einem Neuron liegen (beliebig viele) Eingaben an.
3. Jedes Neuron hat einen individuellen Schwellenwert. Der Schwellenwert muss getroffen oder übertroffen werden, damit ein Neuron eine “1” feuert.
4. Hemmung ist “absolut“: Liegt ein hemmendes Signal zusammen mit nicht-hemmenden Signalen an, feuert das Neuron nicht¹⁰.

tremors of Parkinson's disease, McCulloch thought, could be explained by closed loops of activity connecting the spinal cord and the contracting muscles.“ [Pic04, 178]. *Arbib* vermerkt in [Arb00, 3]: “As another stage in McCulloch's search for the logic of the nervous system, he began thinking about loops in the nervous system. [...] an idea that today we take for granted.“ *Arbib* verweist auf *Lawrence Kubie*, der im Jahr 1930 [Kub30] veröffentlicht, eine Arbeit über “the possible importance of reverberating loops“ [Arb00, 5], die für die Forschungen McCullochs wichtig gewesen ist (vgl. [Arb00, 5])

⁷ Die übernommene Notation gilt als sperrig [AR88, 16]. Eine ausführliche Übersicht der in der PM gebräuchlichen Notation findet sich unter <https://plato.stanford.edu/entries/pm-notation/> (abgerufen 10.8.2023)

⁸ vgl. [Fau94, 26 f.]. Für unsere einführende Betrachtung folgen wir [Roj93, 33 f.] und ignorieren zunächst die Zeitwerte aus der Original-Arbeit: Dort wird für die Signalübertragung ein diskreter Zeitwert $t \in \mathbb{Z}$ vereinbart: Die Übertragung eines Signals dauert eine Zeiteinheit, also $t+1$. Vgl. hierzu Abbildung 3.1.

⁹ Die ausbleibende Exozytose in chemischen Synapsen kann in dem Sinne einer Modellierung als “0“-Signal verstanden werden (siehe Kapitel 2.3.1).

¹⁰ Für eine weitere Ausführung absoluter und relativer Hemmung siehe [Roj93, 42 f.].

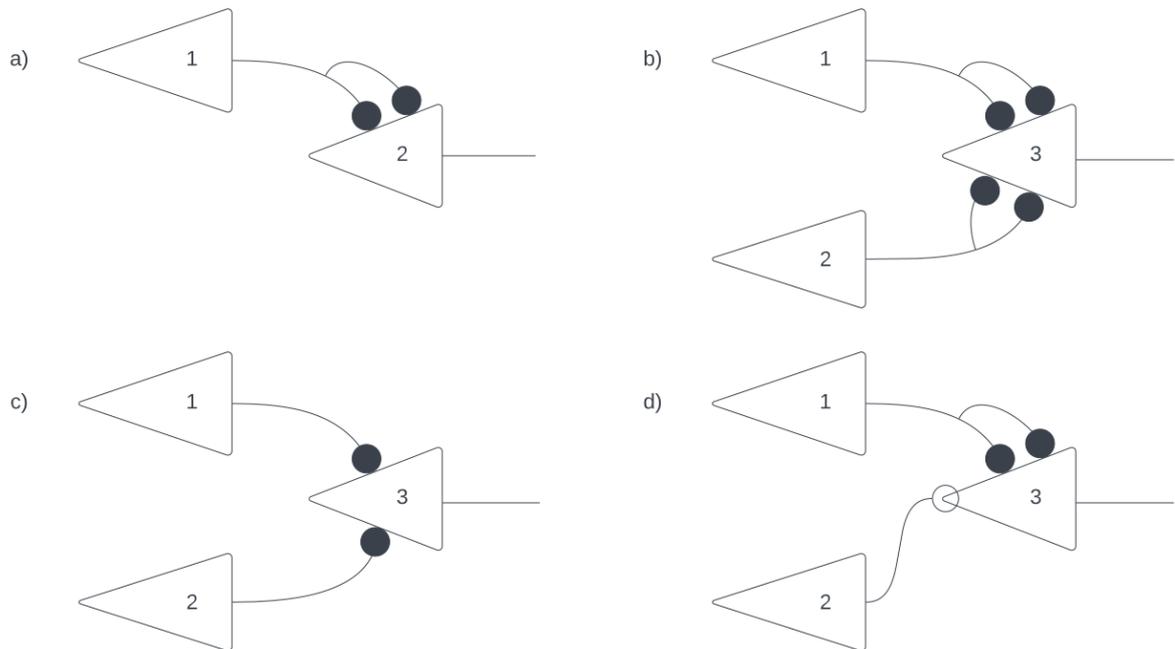


Abb. 3.1. Schematische Darstellung von MCP-Zellen (Quelle: in Anlehnung an [MP43, 105, Figure 1])

Schwarze Kreise sind erregende Verbindungen, offene Kreise hemmende. Im ursprünglichen Modell benötigt ein Neuron zwei erregende Eingaben, um aktiviert zu werden.

In a) ist damit ein Netz dargestellt von zwei Neuronen, bei denen N_2 feuert, wenn N_1 feuert. Unter Berücksichtigung der Zeiteinheit folgt die formale Darstellung $N_2(t) \equiv N_1(t-1)$.

Gleicherweise ergibt sich für b), das N_3 nur feuert, wenn N_1 **oder** N_2 feuern: $N_3(t) \equiv N_1(t-1) \vee N_2(t-1)$. Analog folgt für c) $N_3(t) \equiv N_1(t-1) \wedge N_2(t-2)$.

Für d) ergibt sich somit $N_3(t) \equiv N_1(t-1) \wedge \neg N_2(t-1)$

3.1.3 Aktivierungs- und Eingabefunktion

Mit $n \in \mathbb{N}_0$ und $m \in \mathbb{N}_0$ legen wir fest, dass ein Neuron $n+m$ Eingaben haben soll, mit $n \geq 1 \vee m \geq 1$, wobei n die Anzahl der erregenden und m die der hemmenden Eingaben ist.

Für die Schwellenwertfunktionen eines Neurons N ergeben sich folgende Anforderungen: Die Summe der erregenden Eingabesignale x_1, x_2, \dots, x_n und der hemmenden Eingabesignale y_1, y_2, \dots, y_m muss den für das Neuron festgelegten Schwellenwert $\Theta \in \mathbb{N}_0$ überschreiten, damit als Ausgabe eine 1 erzeugt wird, ansonsten liefert die Funktion eine 0 zurück.

Eingabefunktion

Zur Integration der eingehenden Signale benötigen wir eine **Eingabefunktion**. Der Wert dieser Eingabefunktion wird dann auf eine Funktion angewendet, die entscheidet, ob das Neuron feuert oder nicht - also eine 1 oder 0 ausgibt.

Für die Eingabesignale X

$$X \in \{1, 0\}^{n+m} := (x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_m) \quad (3.1)$$

definieren wir die *Gewichte* $w_+ \in \{2, 1\}$ mit $w_+^1 := 1, w_+^2 := 2$ für erregende Signale, $w_- := -1$ für hemmende Signale (vgl. [Fau94, 27-28]).

Die **Eingabefunktion** g

$$g : \{1, 0\}^{n+m} \rightarrow \mathbb{Z}, X \mapsto \sum_{j=1}^n x_j w_+ + \sum_{k=1}^m y_k w_- \quad (3.2)$$

liefert die Summe der hemmenden und erregenden Eingabesignale zurück.

Aktivierungsfunktion

Die Schwellenwertfunktion wird im Kontext von künstlichen neuronalen Netzen auch **Aktivierungsfunktion** genannt (vgl. [RN12, 847]), da sie entscheidet, ob einzelne Neuronen aktiviert werden oder nicht. In diesem Fall realisieren wir sie als **Treppenfunktion**, die 1 zurückliefert, falls $g(X) \geq \Theta$, und 0 sonst :

$$f : \mathbb{Z} \rightarrow \{0, 1\}, f(g(X)) = f(u) = \begin{cases} 1 & \text{falls } u \geq \Theta \\ 0 & \text{falls } u < \Theta \end{cases} \quad (3.3)$$

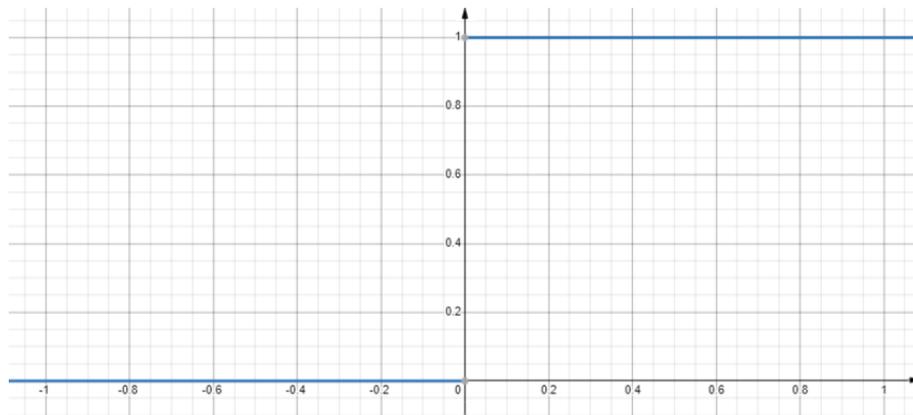


Abb. 3.2. Plot einer Treppenfunktion, hier als Heaviside-Funktion mit $H(x) = \{x < 0 : 0, x \geq 0 : 1\}$. (Quelle: Eigene Darstellung)

Da die Erregung absolut ist, muss Θ die Ungleichung

$$\Theta > \left(\sum_{j=1}^{k w_+^2} 2 \right) + \left(\sum_{j=1}^{k w_+^1} 1 \right) + w_- \quad (3.4)$$

erfüllen, was wir abkürzen können zu

$$\Theta > 2k^{w_+^2} + k^{w_+^1} - 1 \quad (3.5)$$

3.1.4 McCulloch-Pitts-Netz als Graph

Ein zyklensfreies Netz aus MCP-Zellen können wir wie folgt definieren (vgl. [Roj93, 32 “Definition 2.1“]):

Definition 3.1. Ein MCP-Netz ist ein gerichteter Graph¹¹ $G = (V, E)$ mit der Knotenmenge

$$V(G) = \{N_1, N_2, \dots, N_z\} \quad (3.6)$$

und der paarweise verschiedenen Kantenmenge

$$E(G) = \{(N_i, N_j) | i < j\} \quad (3.7)$$

sowie der Kantengewichtsfunktion¹²

$$\gamma : N^2 \rightarrow \{w_+, w_-\}, \gamma(N_i, N_j) = \begin{cases} w_- & \text{falls } (N_i, N_j) \text{ hemmend} \\ w_+ & \text{sonst} \end{cases} \quad (3.8)$$

3.1.5 Implementierung von Booleschen Funktionen

Mit den eingeführten Formalismen können wir nun aussagenlogische Funktionen auf Basis von MCP-Zellen modellieren.

AND (Konjunktion)

In der zweiwertigen Aussagenlogik liefert die **AND**-Funktion (\wedge) nur dann die Ausgabe “wahr“, wenn die darin verknüpften Aussagen jeweils den Wahrheitswert “wahr“ besitzen.

Für “wahr“ = 1 und “falsch“ = 0 sieht die Wahrheitstabelle für zwei Aussagen wie folgt aus (Tabelle 3.1):

¹¹ Gerichtete Kanten eines MCP-Netzes nehmen eigentlich keine Gewichtung der Information vor [Roj93, 40] (*Rojas* verweist darauf, dass Gewichte in absolut hemmenden Leitungen unsinnig sind: Siehe [Roj93, 42]). Sowohl *Minsky* in [Min67, 34] als auch *Rojas* in [Roj93, 32] nutzen unter Berücksichtigung der *absoluten Hemmung* nur die Summe der erregenden Signale und vergleichen diese mit Θ - dieser Vergleich findet nur statt, wenn *kein* hemmendes Signal ankommt: Die Zelle liefert sonst direkt 0 als Ausgabe zurück. *Fausett* nutzt unter Berücksichtigung von Gleichung 3.3 gewichtete Kanten, in ähnlicher Weise auch *Beale und Jackson* in [BJ90, 41]. *Minsky* weist auf die Äquivalenz eines solchen Modells hin (vgl. [Min67, 34 f.]), dessen wir uns im Folgenden auch bedienen werden.

¹² der Einfachheit halber wird bei der Fallunterscheidung keine weitere Unterteilung der für w_+ möglichen Werte $\{1, 2\}$ unternommen

A	B	$A \wedge B$
1	1	1
1	0	0
0	1	0
0	0	0

Tabelle 3.1. Wahrheitstabelle für **AND**

Wir legen für N zwei erregende Eingabesignale x_1, x_2 sowie den Schwellenwert $\Theta = 2$ fest. Die Gewichte für x_1 und x_2 sind jeweils w_+^1 (s. Kantengewichtsfunktion in Definition 3.1).

Mit den oben eingeführten Definition verhält sich unser MCP-Neuron wie folgt (siehe Tabelle 3.2):

$x_1 w_+^1$	$x_2 w_+^1$	$g_{and}(X)$	$f_{and}(x)$
1	1	2	1
1	0	1	0
0	1	1	0
0	0	0	0

Tabelle 3.2. Werte der Aktivierungs- und Eingabefunktion für eine **AND-MCP-Zelle**

OR (Disjunktion)

Für die **OR**-Funktion (\vee) legen wir $\Theta = 2$ und zwei erregende Eingabesignale x_1, x_2 fest. Die Gewichte für x_1 und x_2 sind jeweils w_+^2 : Zusammen mit $\Theta = 2$ reicht also eine erregende Eingabe aus (siehe Tabelle 3.3 sowie Abbildung 3.3):

A	B	$A \vee B$	$x_1 w_+^2$	$x_2 w_+^2$	$g_{or}(X)$	$f_{or}(x)$
1	1	1	2	2	4	1
1	0	1	2	0	2	1
0	1	1	0	2	2	1
0	0	0	0	0	0	0

Tabelle 3.3. Werte der Aktivierungs- und Eingabefunktion für eine **OR-MCP-Zelle**

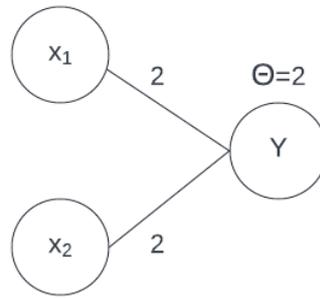


Abb. 3.3. Ein MCP-Neuron zur Modellierung von **OR** (Quelle: Eigene Darstellung)

NOT (Negation)

Für die **NOT**-Funktion (\neg) legen wir $\Theta = 0$ und ein erregendes Eingangssignal x_1 fest. Das Gewicht für x_1 ist w_- : Zusammen mit $\Theta = 0$ würde also ein aktiv anliegendes Signal zu $1w_- = -1 < \Theta = 0$. Durch die resultierende Hemmung kann das Signal die Zelle nicht aktivieren (siehe Tabelle 3.4):

A	$\neg A$	x_1	$g_{\neg}(X)$	$f_{\neg}(x)$
1	0	-1	-1	0
0	1	0	0	1

Tabelle 3.4. Werte der Aktivierungs- und Eingabefunktion für eine **NOT-MCP-Zelle**

NAND (NOT AND)

$\{\neg, \wedge\}$ ¹³ bildet die Operatorenmenge für **NAND**, was als MCP-Netz wie folgt modelliert werden kann (siehe Tabelle 3.5 sowie Abbildung 3.4):

Wir legen $\Theta_1 = 2$ fest (s. **AND**-Zelle), an dem zwei erregende Eingaben anliegen. Θ_2 ist 0 (s. **NOT**-Zelle) und besitzt eine hemmende Eingabe.

N_1 (**AND**) wird aktiviert, wenn sowohl x_1 als auch x_2 ein erregendes Signal weiterleiten. Das Ausgabesignal von N_1 wird dann "invertiert" (siehe Abbildung 3.4).

¹³ $\{\neg, \vee, \wedge\}$ bilden ein vollständiges Operatorensystem, d.h., jede boolesche Funktion lässt sich durch einen aussagenlogischen Ausdruck beschreiben, in dem ausschließlich Operatoren aus dieser Menge vorkommen (vgl. [Hof22, 89]). Da MCP-Zellen die Operatoren \neg, \vee, \wedge darstellen können, ist ein MCP-Netz nach Definition in der Lage, jede boolesche Funktion zu modellieren - hierzu reicht allerdings bereits $\{\neg, \wedge\}$ aus, wie *Sheffer* in [She13] zeigt.

A	B	$\neg(A \wedge B)$
1	1	0
1	0	1
0	1	1
0	0	1

Tabelle 3.5. Die Wahrheitstabelle für **NAND**

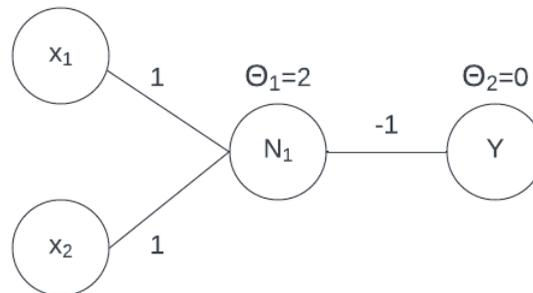


Abb. 3.4. Ein MCP-Netz zur Modellierung von **NAND** (Quelle: Eigene Darstellung)

XOR (exclusive or)

Eine weitere boolesche-Funktion, die aus mehreren MCP-Zellen als Netz modelliert werden kann, ist die **XOR**-Funktion (“exclusive or”). Die Wahrheitstabelle hierfür ist in Tabelle 3.6 zu finden.

A	B	$A \oplus B$
1	1	0
1	0	1
0	1	1
0	0	0

Tabelle 3.6. Die Wahrheitstabelle für **XOR**

Für die Konstruktion des Netzes rufen wir uns Abbildung 3.3 ins Gedächtnis. $\Theta = 2$ wird übertroffen, wenn entweder x_1 oder x_2 ein erregendes Signal weiterleitet. Allerdings wird Θ auch übertroffen, wenn jeweils x_1 und x_2 gleichzeitig aktiv sind, denn dann ist $f_{xor}(g(X)) = 4$. Also muss eine Hemmung zwischengeschaltet

werden (s. Abbildung 3.5): Ein aktives x_1 hemmt N_2 . N_2 leitet in dem Fall 0 an Y weiter. Die Zellen x_2 und N_1 werden analog verbunden.

Die **disjunktive Normalform** (Gleichung 3.9) und die **konjunktive Normalform** (Gleichung 3.10) von **XOR** liefern die Formel für das Netz. Insbesondere die disjunktive Normalform ist in Abbildung 3.5 erkennbar.

$$A \oplus B \equiv (\neg A \wedge B) \vee (A \wedge \neg B) \quad (3.9)$$

$$A \oplus B \equiv (A \vee B) \wedge (\neg A \vee \neg B) \quad (3.10)$$

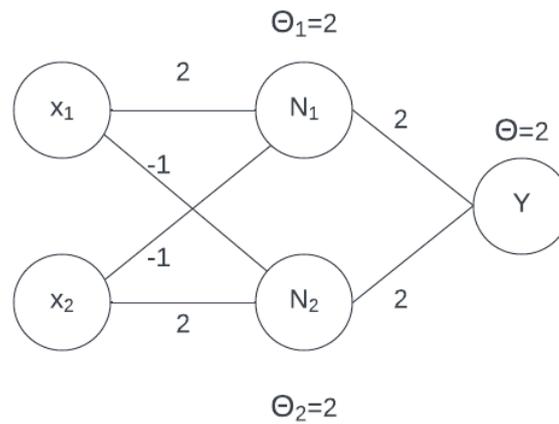


Abb. 3.5. Entwurf für ein MCP-Netz zur Modellierung von **XOR** (Quelle: Eigene Darstellung)

Für weitere Beispiele zu Modellierungen von MCP-Netzen sei auf Anhang C verwiesen.

3.2 Das Rosenblatt-Perzeptron

In Abschnitt 3.1 haben wir gesehen, wie MCP-Netze durch vorhergehende Analyse der Anforderungen und entsprechender Anpassung der Topologie zur Lösung verschiedener Aufgaben imstande sind. *Donald Hebb's* Theorien über *synaptische Plastizität*¹⁴ führen einige Jahre nach dem MCP-Neuron zu einem Modell, das in der Lage ist, sich selbst anzupassen.

3.2.1 Das Perzeptron - ein linearer Klassifizierer

Bereits 1954 wurden Versuche unternommen, lernfähige neuronale Netze zu modellieren (vgl. [Ros62, 24])¹⁵. 1958 schafft es ein Modell, für Aufsehen zu sorgen¹⁶: Das **Perceptron** (im folgenden Perzeptron) (vgl. [AR88, 89])¹⁷. 1957 beschreibt es sein Schöpfer *Frank Rosenblatt* (1928 - 1971) in [Ros57]¹⁸ als Teil eines internen Forschungsprojektes des *Cornell Aeronautical Labors*¹⁹.

3.2.2 Das Modell

Rosenblatt definiert das Perzeptron wie folgt²⁰:

A perceptron is a network of S, A, and R units with a variable interaction matrix V which depends on the sequence of past activity states of the network. ([Ros62, 83 “DEFINITION 17“; Hervorhebung i.O.]

Die darin erwähnten **A**-Units werden definiert mit:

A simple A-unit is a logical decision element, which generates an output signal if the algebraic sum of its input signals, α_i , is equal or greater than a threshold quantity, $\Theta > 0$. The output signal a_i^* is equal to +1 if $\alpha_i \geq \Theta$ and 0 otherwise. If $a_i^* = +1$, the unit is said to be active. ([Ros62, 81 “DEFINITION 9“; Hervorhebung i.O.]

¹⁴ siehe Anhang B.2

¹⁵ Rosenblatt verweist hier auf [FC54]

¹⁶ nicht nur im Wissenschaftsbetrieb, auch in der Öffentlichkeit. In der Presse publikumswirksam, aber wenig vertrauenserweckend: “Frankenstein Monster Designed by Navy Robot That Thinks“ [Ros62, v]. Vgl. dazu aktuelle Schlagzeilen der hiesigen Boulevardpresse zum Thema KI. bild.de titelt: “Übernehmen Computer die Weltherrschaft?“ (<https://www.bild.de/news/ausland/news-ausland/experten-warnen-ki-so-gefaehrlich-wie-pandemien-und-atomkrieg-84130180.bild.html>, abgerufen 27.08.2023) als Reaktion auf “Statement on AI Risk“ des *Center for AI Safety*, in dem zahlreiche Wissenschaftler zur Vorsicht beim Umgang, Einsatz und Forschung von KI mahnen. Im Wortlaut: “Mitigating the risk of extinction from AI should be a global priority alongside other societal-scale risks such as pandemics and nuclear war.“ (<https://www.safe.ai/statement-on-ai-risk>, abgerufen 27.08.2023)

¹⁷ s. auch “Interest in connectionist networks revived dramatically in 1962 with the publication of Frank Rosenblatt’s book *Principles of Neurodynamics*, [...]“ [MP88, xi; Hervorhebung i.O.]

¹⁸ siehe Anhang A.12

¹⁹ Die Forschungseinrichtung gehörte von 1946 bis 1972 zu der Cornell Universität (<https://www.cornell.edu>; ab 1972 gehörte die Forschungseinrichtung dann zu der Calspan Corporation, <https://calspan.com>, beides abgerufen 18.08.2023). Rosenblatt macht an der Cornell Universität 1950 seinen A.B, 1956 seinen Ph.D. und wird dort bis zu seinem Lebensende (1971) als Psychologe und Neurobiologe forschen und lehren.

²⁰ Ausführliche Definitionen aller Zustände, Signale und Funktionen in [Ros62, 79 - 94]

Ähnlichkeiten zwischen den A-Units und zu der in Abschnitt 3.1.3 eingeführten Aktivierungsfunktion sind durchaus erkennbar: Rosenblatt selber weist darauf hin, dass er sein Modell direkt von dem von McCulloch und Pitts eingeführten Modell ableitet²¹. Darüber hinaus weist er auch auf Einflüsse von Hebb und von Neumann hin (siehe [Ros62, 5]).

Rojas schreibt, dass das klassische Rosenblatt-Perzeptron in einem Netz von Eingabe- und Ausgabeknoten gewichtete Verbindungen nutzt - die Knoten selber sind Schwellenwertelemente, Verbindungen werden stochastisch ermittelt (siehe [Roj93, 51]). Er weist ebenda darauf hin, dass das Modell nach Rosenblatts Veröffentlichung analysiert und verfeinert wurde, u. a. von *Minsky und Papert* in [MP88], die wie folgt definieren²²:

A perceptron is a device capable of computing all predicates which are linear in some given set Φ of partial predicates. ([MP88, 12; Hervorhebung i.O.]

Prädikate sind hier Verbindungen zu den Eingabesignalen, die einen Wahrheitswert 0 oder 1 basierend auf der Eingabe X berechnen. Die Ausgaben der Prädikate werden individuell gewichtet und an die Zelle weitergeleitet, die die Aktivierungsfunktion implementiert (vgl. [Roj93, 52 f.] und [MP88, 8-12]).

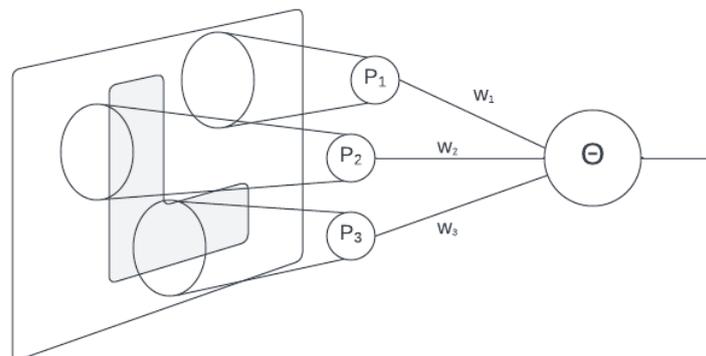


Abb. 3.6. Rosenblatt-Perzeptron. (Quelle: in Anlehnung an [Roj93, 53, Abb. 3.2])

Schematische Darstellung eines Perzeptron-“Automats“ (siehe [Ros57, 7]): Eingaben werden an die Prädikate P_n weitergeleitet, deren binäre Ausgaben mit den Gewichten w_n multipliziert werden. Die Summe der Produkte wird mit dem Schwellenwert Θ verglichen (vgl. [Roj93, 53]).

Die Eingabefunktion setzt sich dann wie Gleichung 3.2 aus der Summe der Produkte der Prädikate $P_i \in \Phi$ (wobei $P_i(X) \in \{0, 1\}$) und den Gewichten $w_i \in \mathbb{R}$

²¹ vgl. auch “Ein einfaches Perzeptron ist eine McCulloch-Pitts-Zelle, die ihre Eingabe gewichtet berechnet.“ in [Roj93, 57, “Definition 3.1“; Hervorhebung i.O.]

²² wir werden später sehen, dass die Eigenschaft “linear“ in dieser Definition von besonderer Bedeutung sein wird.

der Verbindungen zusammen (s. Gleichung 3.11), und die Aktivierungsfunktion (s. Gleichung 3.12) ist wieder eine Treppenfunktion mit dem reellen Schwellenwert Θ :

$$g := g(X) = \sum_{i=1}^n P_i(X)w_i \quad (3.11)$$

$$f := f(g(X)) = f(x) = \begin{cases} 1 & \text{falls } x \geq \Theta \\ 0 & \text{falls } x < \Theta \end{cases} \quad (3.12)$$

3.2.3 Lineare Trennbarkeit

Mit Gleichung 3.12 folgt für eine Eingabe, dass sie entweder eine 0 oder 1 als Ausgabe erzeugt. Es liegt ein binärer Wertebereich vor, der als zwei unterschiedliche Klassen aufgefasst werden kann: Eingabedaten können entsprechend ihrer Ausgabe einer der beiden Klassen zugeordnet werden (vgl. [RN12, 812]). Im Folgenden wollen wir die Zusammenhänge geometrisch darstellen. Der Einfachheit halber beschränken wir uns hierzu auf den zweidimensionalen Raum \mathbb{R}^2 und betrachten dort die 1. Winkelhalbierende im I. und III. Quadranten des kartesischen Koordinatensystems. Die zugehörige Gerade L ist

$$L = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 = x_2\} \quad (3.13)$$

Für beliebige Punkte $(x_1, x_2) \in \mathbb{R}^2$ gilt damit offensichtlich

$$x_1 - x_2 \begin{cases} > 0 & \text{falls } x_1 > x_2 \\ = 0 & \text{falls } x_1 = x_2 \\ < 0 & \text{falls } x_1 < x_2 \end{cases} \quad (3.14)$$

Die Gerade L repräsentiert eine *Hyperebene*²³ im \mathbb{R}^2 . Punkte, die nicht zu dieser Hyperebene gehören, liegen in zwei unterschiedlichen *Halbräumen*²⁴ (siehe geometrische Darstellung in Abbildung 3.8).

Wir können feststellen, dass

- Punkte, die $x_1 - x_2 > 0$ erfüllen (im folgenden M_-) in dem Halbraum *unter* der durch L beschriebenen Gerade liegen
- Punkte, für die $x_1 - x_2 < 0$ gilt (im folgenden M_+) *über* der durch L beschriebenen Gerade liegen
- Punkte mit $x_1 - x_2 = 0$ *auf* der Geraden liegen (im folgenden als Teilmenge von M_- ²⁵)

M_+ und M_- sind damit *linear separierbar*. Nach *Rojas* lautet die Definition für **Lineare Trennbarkeit**:

²³ siehe [Bur+12, 81, Definition 2.3]

²⁴ Renze, John; Uznanski, Dan; and Weisstein, Eric W. "Half-Plane." From MathWorld—A Wolfram Web Resource. <https://mathworld.wolfram.com/Half-Plane.html> (abgerufen 21.08.2023)

²⁵ Wenn die Hyperebene selbst im Halbraum enthalten ist, spricht man von einem *abgeschlossenen Halbraum*. (<https://de.wikipedia.org/wiki/Halbraum>, abgerufen 22.08.2023)

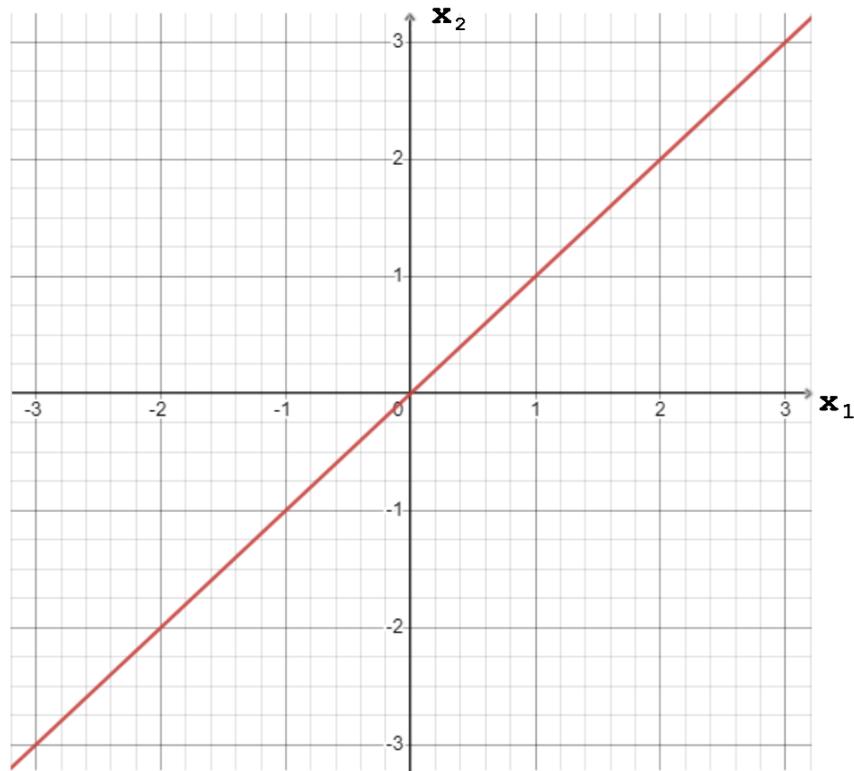


Abb. 3.7. 1. Winkelhalbierende im kartesischen Koordinatensystem (Quelle: Eigene Darstellung)

Zwei Mengen A und B von Punkten in einem n -dimensionalen Raum sind *linear trennbar*, falls $n+1$ reelle Zahlen w_1, \dots, w_{n+1} existieren, so dass für jeden Punkt $x_1, \dots, x_n \in A$ gilt

$$\sum_{i=1}^n w_i x_i \geq w_{n+1} \quad (3.15)$$

und für jeden Punkt $x_1, \dots, x_n \in B$

$$\sum_{i=1}^n w_i x_i < w_{n+1} \quad (3.16)$$

([Roj93, 60 f., “Definition 3.2“; Hervorhebung i.O.; Nummerierung eigene])

Als Beispiel sei $M_- = \{(x_1, x_2) \in \mathbb{R}_{0+}^2 : x_1 \geq x_2\}$ und $M_+ = \{(x_1, x_2) \in \mathbb{R}_{0+}^2 : x_1 < x_2\}$. Weiter sei $w_1 = -1, w_2 = 1, w_3 = 0$. Die lineare Trennbarkeit der beiden Mengen lässt sich leicht nachweisen²⁶:

²⁶ entspricht zwei Punkt Mengen, die durch die 1. Winkelhalbierende im I. Quadranten des kartesischen Koordinatensystems separierbar sind

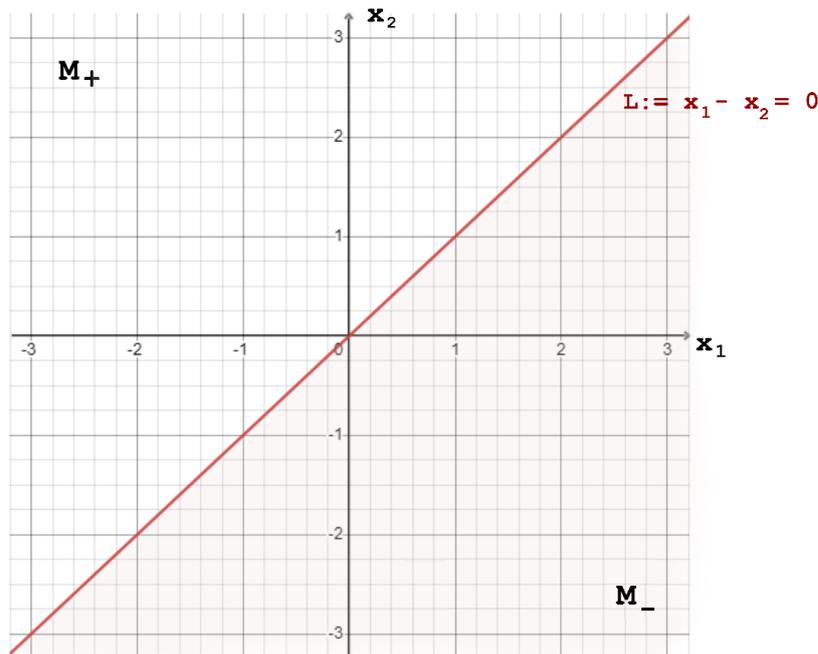


Abb. 3.8. Halbräume im \mathbb{R}^2 (Quelle: Eigene Darstellung)

Skizzierung der durch die 1. Winkelhalbierende entstandenen Halbräume. Die Mengen M_- und M_+ sind linear separierbar, die Gleichung für die *Trenngerade* hierzu lautet $x_1 - x_2 = 0$

1. Fall $x_1 = x_2$: Es gilt $w_1x_1 + w_2x_2 = w_1x_1 + w_2x_1 = -x_1 + x_1 = w_3 = 0$. Mit $0 \geq 0$ ist somit Gleichung 3.15 erfüllt.
2. Fall $x_1 < x_2$: Es gilt $w_1x_1 = -x_1$. Addition von w_1x_1 auf beiden Seiten von $x_2 > x_1$ liefert $x_2 + (-x_1) = w_2x_2 + w_1x_1 > 0 = w_3$ und erfüllt Gleichung 3.15.
3. Fall $x_1 > x_2$: Es gilt wieder $w_1x_1 = -x_1$. Addition auf beiden Seiten von $x_1 > x_2$ liefert $w_3 = 0 > x_2 + (-x_1) = w_2x_2 + w_1x_1$ und erfüllt Gleichung 3.16.

Insgesamt kann ein Perzeptron als Funktion verstanden werden, was sich auch in der Definition nach *Ertel* widerspiegelt:

Sei $w = (w_1, \dots, w_n) \in \mathbb{R}^n$ ein Gewichtsvektor und $x \in \mathbb{R}^n$ ein Eingabevektor. Ein **Perzeptron** stellt eine Funktion $P : \mathbb{R}^n \rightarrow \{0, 1\}$ mit

$$P(x) = \begin{cases} 1 & \text{falls } wx = \sum_{i=1}^n w_i x_i > 0 \\ 0 & \text{sonst} \end{cases}$$

dar. ([Ert21b, 212, "Definition 8.3"; Hervorhebung i.O.]

3.2.4 Die Lernregel

Die Eigenschaft linearer Trennbarkeit von Daten ist eine wesentliche Voraussetzung dafür, dass ein Perzeptron **konvergiert**: Die **Lernregel** des Perzeptrons passt während der Laufzeit die Gewichte $w_1 \dots w_n$ solange an, bis sie - eingesetzt in eine

lineare Gleichung (vgl. [Ert21c, 311]) - die n -dimensionalen Daten entsprechend Gleichung 3.12 *klassifizieren* kann. Aus diesem Grund wird das Perzeptron auch **linearer Klassifizierer** genannt (vgl. [Ert21b, 210-216]).

Das Perzeptron **lernt** diese Gewichte zunächst durch *Trainingsdaten*²⁷. Jeder Eintrag dieser Trainingsdaten ist einer erwarteten Ausgabe zugeordnet. Der Algorithmus besteht aus folgenden Schritten (vgl. [RM87, 65] sowie [RN12, 842])²⁸:

1. Wähle einen Datensatz und berechne die Ausgabe.
2. Wenn die Ausgabe 1 ist, obwohl sie 0 sein sollte (Fehler²⁹ = -1), verringere die Gewichte.
3. Wenn die Ausgabe 0 ist, obwohl sie 1 sein sollte (Fehler = 1), erhöhe die Gewichte.
4. Wenn die Ausgabe korrekt ist, passe die Gewichte nicht an.

Die Schritte werden so lange durchlaufen, bis für alle Trainingsdaten die Ausgabe korrekt ist, oder eine maximale Anzahl von Trainingsläufen erreicht wurde. Einen Trainingslauf nennt man dabei **Epoche**³⁰. Sind die Trainingsdaten linear separierbar, konvergiert³¹ das Perzeptron nach einer endlichen Zahl von Epochen (vgl. [MP88, 164])³², und ist danach in der Lage weitere Daten zu **generalisieren** (vgl. [Ert21b, 202]).

Wir betrachten jetzt eine Möglichkeit, die y -Koordinate der Trenngerade anzupassen, falls Punkte vorliegen, die nicht durch eine Ursprungsgerade separierbar sind (siehe Abbildung 3.9).

Dies erreicht man mit einer sogenannten **bias unit**. Das Bias-Gewicht³³ ist ein Wert, der zu der Gleichung 3.11 hinzuaddiert wird. In der geometrischen Darstellung sorgt dieser Wert für eine Verschiebung der Ursprungsgeraden entlang der x_2 -Achse³⁴.

Den Eingabedaten wird ein fester Eingabewert $x_{n+1} = 1$ hinzugefügt: Der Eingabevektor $x \in \mathbb{R}^n$ wird *erweitert*: $(x_1, \dots, x_n, 1)$ (vgl. [Roj93, 58]). Der bias b wird für unser Beispiel im \mathbb{R}^2 mit Gleichung 3.11 in die Berechnung des Schwellenwerts miteinbezogen:

²⁷ “supervised learning“: überwachtes lernen; vgl. [RN12, 811] sowie [Fau94, 15]. *Arbib et al.* weisen darauf hin, dass *überwachtes Lernen* mit dem Perzeptron eingeführt wurde: “Supervised learning adjusts the weights in an attempt to respond to explicit error signals provided by a ‘teacher,’ which may be external, or another network in the same ‘brain’. This model was introduced in the perceptron model“ [Arb03, 30]

²⁸ Die Beziehung zu der Hebbschen Lernregel formulieren *Arbib et al.*: “The best-known perceptron learning rule strengthens an active synapse if the efferent neuron fails to fire when it should have fired, and weakens an active synapse if the neuron fires when it should not have done so.“ [Arb03, 20]. Vgl. hierzu auch *Rosenblatt*: “[a perceptron has] a tendency to develop ‘cell assemblies’ (in Hebb’s sense), and these cell-assemblies tend to rival one another for dominance at all times.“ [Ros62, 464]

²⁹ Der **Fehler** ist die Differenz von erwarteter Ausgabe und tatsächlicher Ausgabe.

³⁰ siehe [Fau94, 436]

³¹ “iterative training processes converge if the weight updates reach equilibrium (stop changing)“ [Fau94, 425 “Convergence“].

³² Das Konvergenz-Theorem besagt: “if a linear separation exists, the perceptron error-correction scheme will find it.“ [Arb03, 20] Beweise führen [Ros62, 111 ff.], [MP88, 168 ff.] sowie [Nov62].

³³ vgl. [RN12, 839]

³⁴ im \mathbb{R}^n durch eine Hyperebene im Ursprung (vgl. [Ert21b, 215]).

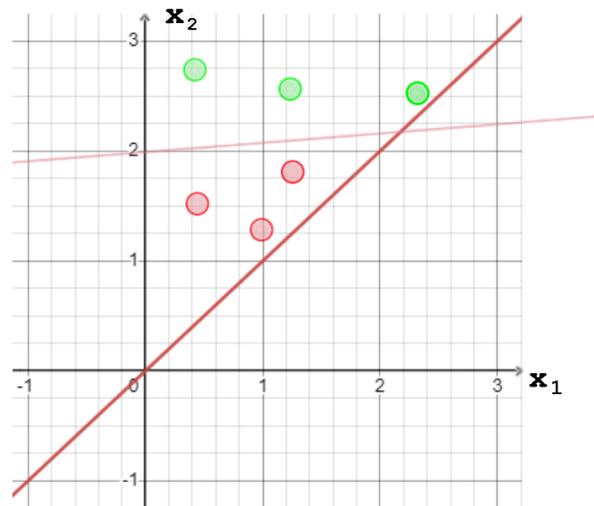


Abb. 3.9. Punkte im \mathbb{R}^2 , die nicht durch eine Ursprungsgerade separierbar sind. Angedeutet eine mögliche Trenngerade, die durch $(0, 2)$ geht. (Quelle: Eigene Darstellung)

$$P(x) = \begin{cases} 1 & \text{falls } wx > 0 \\ 0 & \text{sonst} \end{cases} \quad (3.17)$$

wobei

$$wx = b + \sum_{i=1}^n w_i x_i \quad (3.18)$$

Die Gleichung für die Trenngerade für unser Beispiel im \mathbb{R}^2 lautet somit

$$b + w_1 x_1 + w_2 x_2 = 0 \quad (3.19)$$

Wenn wir b auf die rechte Seite der Gleichung bringen, kann b auch als Schwellenwert $\Theta = -b$ betrachtet werden:

$$w_1 x_1 + w_2 x_2 = \Theta \quad (3.20)$$

Für unser x_2 im \mathbb{R}^2 folgt dann insgesamt mit $w_2 = 1$ und $x_1 = 0$:

$$x_2 = \Theta/w_2 - (w_1/w_2)x_1 = \Theta \quad (3.21)$$

x_2 ergibt sich somit als der horizontale Abstand vom Ursprung.

Mit b als Teil der Eingabe folgt auch dessen Gewichtung $bw_{n+1} = w_{n+1}$. Mit 0 als Schwellenwert wird dadurch eine Verschiebung der Trenngerade entlang der

x_2 -Achse im \mathbb{R}^2 bzw. eine Verschiebung der Hyperebene im \mathbb{R}^n ermöglicht (vgl. [Ert21b, 215])³⁵.

3.2.5 Die XOR-Funktion

Wenn ein Perzeptron nicht konvergiert, kann es ausreichen, die Anzahl der Epochen zu erhöhen, damit ein passender Gewichtsvektor gefunden wird³⁶.

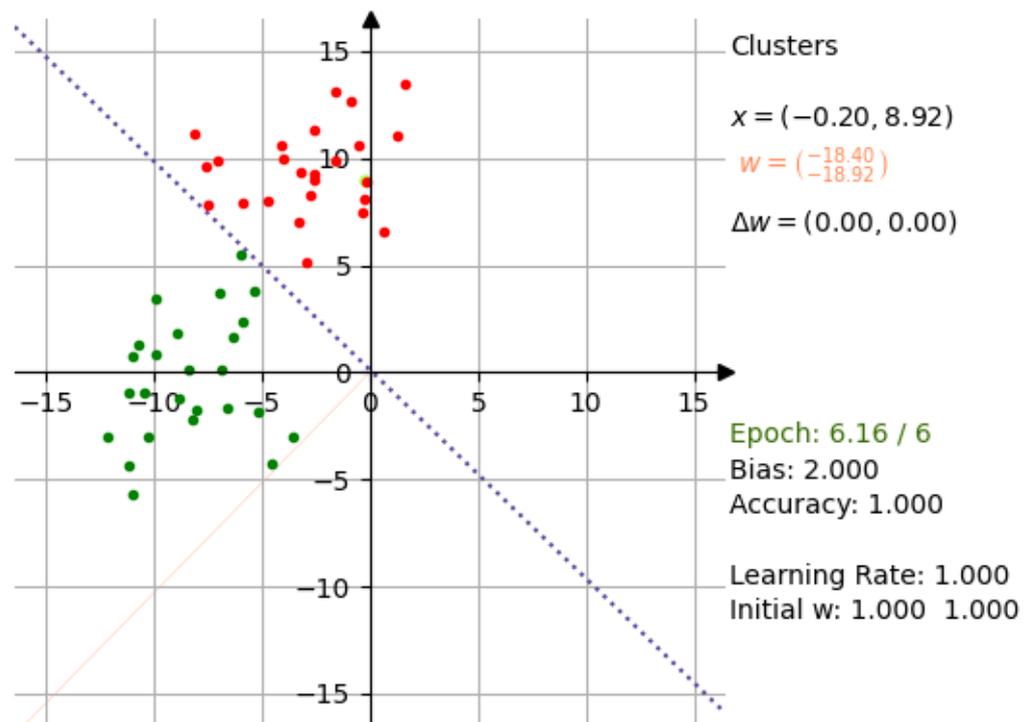


Abb. 3.10. Perzeptron-Training für große Datenmengen. (Quelle: Eigene Darstellung)

Ein Perzeptron wird mit einer großen Datenmengen (50 Einträge) trainiert. Nach knapp 300 Trainingschritten (in der 6ten Epoche) wird die Trenngerade (gestrichelt) gefunden.

³⁵ einen Überblick über die geometrischen Zusammenhänge liefert “Einführung in die Neuroinformatik“ von Prof. Dr. G. Sommer <https://www.informatik.uni-kiel.de/inf/Sommer/doc/Downloads/Skripte/neuroskript.pdf>:33 ff. (abgerufen 27.08.2023)

³⁶ *Arbib et al.* berufen sich auf das Konvergenz-Theorem: “[das Rosenblatt Perzeptron] does not yield an endless seesaw, but will eventually converge to a correct set of weights if one exists, albeit perhaps after many iterations through the set of trial patterns.“ [Arb03, 20]; *Minsky und Papert* formulieren lose: “if the sets are separable [...], then the program will separate them“ [MP88, 165]. Sie stellen im Hinblick auf Parameteranpassungen fest: “Usually, when a failure occurred, neither prolonging the training experiments nor building larger machines helped.“ [MP88, xi]

Allerdings kann es bereits bei wenigen Daten und beliebig großer Epochenzahl vorkommen, dass ein Perzeptron nicht konvergiert, nämlich wenn die Daten nicht linear separierbar sind (vgl. [Arb03, 20]).

Als Beispiel betrachten wir die boolesche Funktion **XOR** (vgl. Tabelle 3.6). In Abbildung 3.11 ist die geometrische Repräsentation der möglichen Interpretationen für $A \oplus B$ dargestellt. Zwar lassen sich die Elemente separieren, aber nicht linear. Es müsste sonst ein w_1, w_2 existieren, das folgende Ungleichungen erfüllt:

$$w_1 0 + w_2 0 < \Theta$$

$$w_1 1 + w_2 0 \geq \Theta \implies w_1 \geq \Theta$$

$$w_1 0 + w_2 1 \geq \Theta \implies w_2 \geq \Theta$$

$$w_1 1 + w_2 1 < \Theta$$

Offensichtlich kann $w_1 + w_2 < \Theta$ nicht erfüllt werden, wenn gleichzeitig $w_1 \geq \Theta$ und $w_2 \geq \Theta$ gilt (s. Abbildung 3.11).

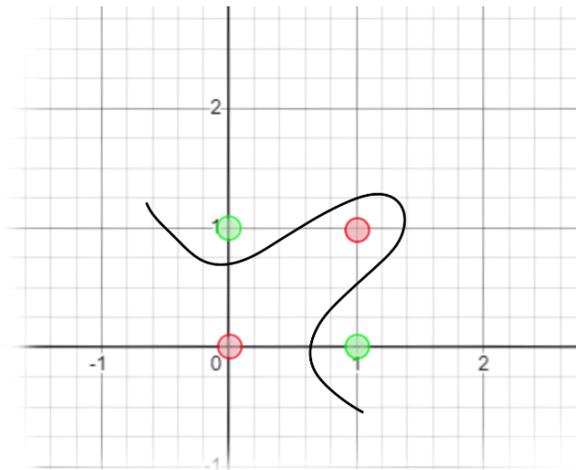


Abb. 3.11. XOR im \mathbb{R}^2 . (Quelle: Eigene Darstellung)

Interpretationen der XOR-Funktion im kartesischen Koordinatensystem. Hier existiert keine Trenngerade für die Punkte $\{(0, 0), (1, 1)\}$ und $\{(0, 1), (1, 0)\}$.

Für weitere Anwendungen des Perzeptron-Algorithmus sei auf Anhang D verwiesen.

Künstliche Neuronale Netze

4.1 Renaissance der neuronalen Netze

In den 1980er Jahren sahen die großen Industrienationen in der Erforschung von KI einen Wettbewerbsvorteil, nachdem die Technologie durch den Einsatz von Expertensystemen (siehe Anhang A.6) Einzug in die Wirtschaft gefunden hatte: So ermöglichte R1/XCON bei dem einsetzenden Unternehmen DEC (Digital Equipment Corporation) Einsparungen in Millionenhöhe¹: Die Domäne von R1 war die regelbasierte Konfiguration von VAX-11/780 Systemen nach Kundenanforderungen². Japan kündigte 1981 das Projekt *5th Generation Computer Systems* (vgl. [Gar19]) an, einen Zehnjahresplan für den Aufbau “intelligenter Computer“ [RN12, 48]³, und in Großbritannien sorgte der Alvey-Report⁴ für eine Wiederaufnahme finanzieller Unterstützung, die durch den Lighthill-Report aufgehoben worden war (vgl. [RN12, 48])^{5,6}.

Auch der technologische Fortschritt begünstigte das Wiederaufleben des Interesses an neuronalen Netzen, wie *Olazaran* in Bezug auf die Modellierung paralleler Prozesse mit Hilfe von neuronalen Netzen anmerkt:

[...] increases in computer power and speed due to parallelism will undoubtedly favour neural net research. ([Ola96, 644])

¹ vgl. [RN12, 48]; ebenda wird es von *Russell und Norvig* als das “erste erfolgreiche kommerzielle Expertensystem“ bezeichnet

² siehe hierzu [McD80] sowie [Hor90, 63]

³ Zusammenfassend war das Ziel des 5th Generation Computer Systems (FGCS)-Projekt: “Its ultimate goal is to develop integrated systems, both hardware and software, suitable for the major computer application in the next decade, identified by the Japanese as ‘knowledge information processing’.“ [Sha83, 637].

⁴ https://www.chilton-computing.org.uk/inf/literature/reports/alvey_report/overview.htm - empfohlene Massnahmen: https://www.chilton-computing.org.uk/inf/literature/reports/alvey_report/p008.htm (beide abgerufen 31.08.2023)

⁵ zu weiteren Hintergründen hierzu siehe Anhang A.8.

⁶ In Deutschland wird 1988 die DFKI GmbH (Deutsches Forschungszentrum für Künstliche Intelligenz) gegründet, eine “wirtschaftsnahe Forschungseinrichtung“ auf “dem Gebiet innovativer Softwaretechnologien auf der Basis von Methoden der Künstlichen Intelligenz“ (https://www.dfki.de/fileadmin/user_upload/DFKI/Medien/Ueber_uns/DFKI_im_Ueberblick/Unternehmensprofil/20210120_DFKI_Unternehmensprofil_DE.pdf, abgerufen 31.08.2023)

Mit den in den 1980er Jahren verfügbaren Supercomputer und Parallelrechner⁷ erhält auch der **Konnektionismus** Auftrieb, der **neuronale Netze** als Grundlage hat (vgl. [Dor91, 15]), und mit dem sich Modelle wieder mehr an den biologischen Vorbildern orientieren sollten⁸.

4.2 Mehrschichtige neuronale Netze

Bislang haben wir überwiegend künstliche Neuronen betrachtet, bei denen die Eingabe direkt mit der Ausgabe verbunden ist. Allerdings haben wir bereits für komplexe Boolesche Funktionen in Abschnitt 3.1.5 gesehen, dass ein Verbund von mehreren MCP-Zellen in der Lage ist, auch Funktionen für nicht linear separierbare Daten zu modellieren. Hierzu hatten wir das MCP-Netz in zwei Schichten aufgeteilt, in denen die erste Schicht $X_1 := (\neg A \wedge B)$ sowie $X_2 := (\neg B \wedge A)$ und die zweite Schicht $X_1 \vee X_2$ formt, was die disjunktive Normalform von XOR ist (Tabelle 3.6).

Bei dem Rosenblatt-Perzeptron, das alleine nicht in der Lage ist, XOR zu erlernen, handelt es sich um ein **Einschichtiges neuronales Feedforward-Netz**. Das bedeutet, dass es nur Eingabe und Ausgabe gibt und die Informationen ausschliesslich in Richtung Ausgang fliessen (vgl. [RN12, 848]).

Ein **mehrschichtiges Perzeptron** (MLP, *multi-layer perceptron*) [GBC18, 6] ist in der Lage, nicht-linear-separierbare Daten zu klassifizieren. Ein MLP repräsentiert ein tiefes Feedforward-Netz, in dem die Eingabeschicht (**Input Layer**) und die Ausgabeschicht (**Output Layer**) über weitere Schichten von Neuronen (**hidden layer**) verbunden ist; die Neuronen in diesen Schichten implementieren Eingabe- und Aktivierungsfunktion und leiten ihre Berechnungen an die nächsten Zellen bis zu der Ausgabeschicht weiter. *Murtagh* zeigt in [Mur91, 184 f.] die Modellierung von XOR anhand eines MLPs.

4.2.1 Backpropagation

In gleicher Arbeit zeigt *Murtagh* ein MLP, das mittels **Backpropagation**⁹ Daten klassifiziert, die keinen linearen Zusammenhang besitzen (siehe [Mur91, 185 f.]). Der Algorithmus geht auf *Rumelhart, Hinton und Williams* zurück, die in [RM87, 318 ff.] eine Methode¹⁰ vorstellen, um berechnete Werte *rückwärts* in das Netz einzuspeisen. Hierbei wird für die Netzausgabe (*forward pass*) ein Approximationsfehler berechnet, der als Basis für die Gewichtsänderungen beim rückwärtigen Lauf (*backward pass*) bis zur ersten verborgenen Schicht genutzt wird. Der Vorgang

⁷ Auch *Fausett* nennt bessere Rechenleistung als einen Grund für den erneuerten Enthusiasmus (vgl. [Fau94, 26]). Im Kontext der in diesem Kapitel vorgestellten mehrschichtigen Netze und ihren versteckten Schichten ist nachvollziehbar, dass mehr Rechenleistung den komplexen Verfahren entgegenkommt. *Goodfellow et al.* gehen auf die Korrelation Modellgrösse und Anzahl der Nervenzellen in einem menschlichen Gehirn in [GBC18, 24 f.] ein

⁸ vgl. [RM87, 43] sowie [GBC18, 18 f.]

⁹ "das meist genutzte neuronale Modell" [Ert21c, 313]

¹⁰ "The ability to create useful new features distinguishes back-propagation from earlier, simpler methods such as the perceptron-convergence procedure." [RHW86, 533]

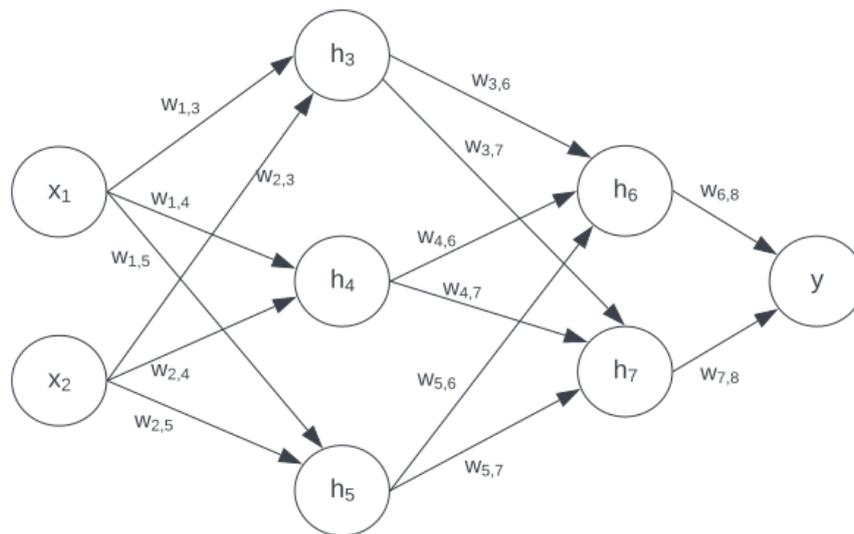


Abb. 4.1. Exemplarische Darstellung der Architektur eines einfachen Feed-Forward-Netzes. (Quelle: Eigene Darstellung)

Ein Feed-Forward-Netz mit zwei Eingabe-einheiten x_1, x_2 , zwei versteckten Schichten und einer Ausgabeneinheit y . Das dargestellte Netz ist vergleichsweise klein. Große Modelle wie Microsoft's Turing-NLG besitzen bis zu 105 Schichten und 530.000.000.000 Parameter (<https://www.microsoft.com/en-us/research/blog/using-deepspeed-and-megatron-to-train-megatron-turing-nlg-530b-the-worlds-largest-and-most-powerful-generative-language-model/>, abgerufen 27.09.2023).

(forward pass, backward pass, forward pass...) wird so lange für alle Trainingsbeispiele wiederholt, bis die Gewichte sich nicht mehr ändern, oder eine andere Schranke (wie bspw. Epoche oder Zeit) erreicht ist (vgl. [Ert21c, 315])¹¹.

Die mathematische Basis für Backpropagation ist das **Gradientenabstiegsverfahren**, das dabei hilft, in einem neuronalen Netz Parameter für möglichst optimale (d.h. kleine) Verlust-Werte (also geringe Fehler-Werte) zu finden (vgl. [RN12, 837]; s. a. Abbildung 4.3)¹². Als Aktivierungsfunktion unterstützt eine **Sigmoid**¹³-Funktion¹⁴ (Gleichung 4.1 sowie Abbildung 4.2) aufgrund ihres nicht-linearen Charakters eine größere Klasse darstellbarer Funktionen und damit Lösungen für

¹¹ ausführlicher Algorithmus in [RN12, 853, Abbildung 18.24].

¹² ein kompakter Überblick zum Thema "Optimierung auf Gradientenbasis" findet sich in [GBC18, 90 ff.]

¹³ *sigmoid*: "Sigma", griechischer Buchstabe Σ , entspricht dem lateinischen "S"; "eidos" (griechisch) Form, Gestalt

¹⁴ "Historisch wurde Backpropagation mit der Sigmoidfunktion implementiert." [Ert21c, 314, Fussnote 4]. Siehe hierzu [RM87, 328 f.]. *Ertel* fügt hinzu: "Mittlerweile haben sich jedoch andere Funktionen als besser bewährt." In [Ert21c, 319 f.] geht der Autor auf "Probleme und Verbesserungen" des mittlerweile über 30 Jahre alten Verfahrens ein. *Goodfellow et al.* stellen fest, dass Backpropagation der "vorherrschende Ansatz für das Training tiefer Modelle" (im Jahr 2018) ist [GBC18, 19].

Probleme, die ein klassisches Perzeptron nicht lösen kann (vgl. [Ert21c, 316])¹⁵ ¹⁶.

Sigmoid:

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (4.1)$$

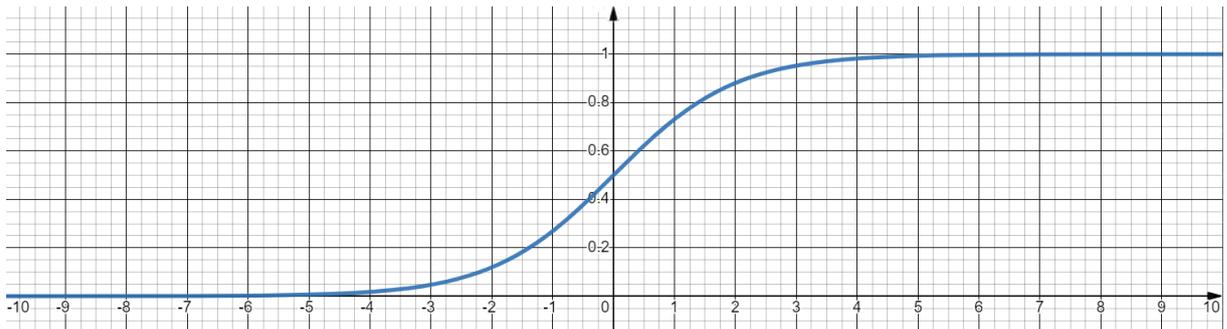


Abb. 4.2. Plot einer Sigmoid-Funktion. (Quelle: Eigene Darstellung)

4.2.2 Hopfield-Netz

Der Physiker *John J. Hopfield* stellt 1982 in [Hop82] ein Modell für ein *asynchrones* neuronales Netz¹⁷ vor, das auf neurobiologischen Aspekten beruht (vgl. [Hop82, 2554]) und 5 Jahre später in Zusammenarbeit mit AT&T Bell Laboratories¹⁸ in Form eines “neural network chip“ [AR88, 457]¹⁹ als Hardware umgesetzt wird: Das Modell ist ein **rekurrentes Netz**²⁰ ohne Schlingen ($w_{ii} = 0$) (vgl. [Ert21c, 291])²¹ mit bidirektionalen symmetrischen Kanten ($w_{ij} = w_{ji}$) und fungiert als **Assoziativspeicher**, also als Netz, das, wenn es für eine Eingabe x die Ausgabe y liefert, dieselbe Ausgabe y auch für x' berechnet, sofern x' nahe genug bei x liegt

¹⁵ In [RM87, 311 ff.] zeigen *Rumelhart, Hinton und Williams* die Architektur eines mehrschichtigen neuronalen Netzes, das das XOR-Problem über Backpropagation löst. In [RHW86, 536] grenzen sie ihr Modell vom biologischen Lernen ab: “The learning procedure, in its current form, is not a plausible model of learning in brains“.

¹⁶ *Ertel* weist darauf hin, dass die Klasse der darstellbaren Funktionen eines Perzeptrons erhöht wird, wenn man eine variable Sigmoid-Funktion verwendet [Ert21c, 316]. Siehe hierzu auch “Sigmoid-Perzeptron“ in [RN12, 847]

¹⁷ in asynchronen Modellen werden die Aktivierungen der Neuronen unabhängig voneinander und zu unterschiedlichen Zeitpunkten berechnet (vgl. [Roj93, 49] sowie ebenda S. 282)

¹⁸ ehemalige Forschungsabteilung der Telefongesellschaft AT & T; gehört seit 2016 zu Nokia (<https://www.bell-labs.com>, abgerufen 03.09.2023)

¹⁹ <https://techmonitor.ai/technology/att-creates-parallel-neural-net-chip-to-solve-routin-g-problems>

²⁰ Ein rekurrentes neuronales Netz “speist seine Ausgaben wieder in seine eigenen Eingaben ein“ [RN12, 847]. Im Gegensatz zu Feedforward-Netzen sind rekurrente Netze also *rückgekoppelt*

²¹ Eine *Schlinge* (engl. *loop*) in einem Graphen ist ein Kantenzug, der einen Knoten mit sich selbst verbindet [Die17, 30]. Ein Beispiel für ein Netz mit Schlingen ist das **MAXNET** (siehe [Lip87]), das zu der Gruppe der **competitive nets** gehört: Es kann als Subnetz zur Ermittlung der Zelle mit dem größten Aktivierungswert genutzt werden (vgl. [Fau94, 158 f.]).

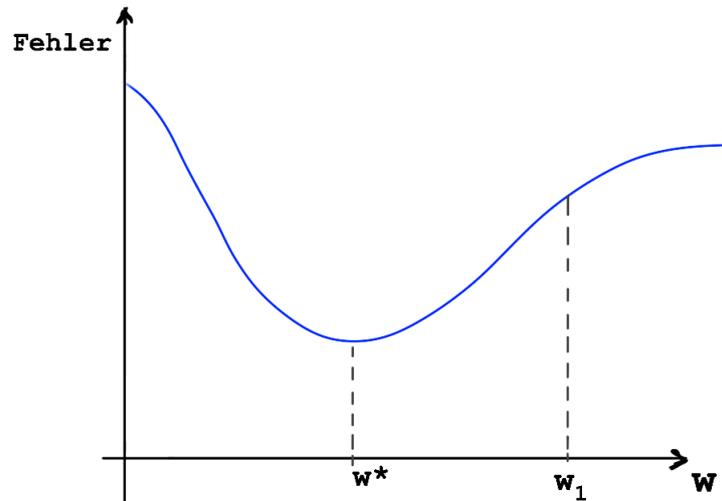


Abb. 4.3. Skizzierung des Zusammenhangs Fehlerfunktion und lokale Minima (Quelle: in Anlehnung an [Son22, 52, Abb. 2.15])

Die y -Achse repräsentiert den Fehlerwert, die x -Achse den berechneten Gewichtsvektor w_1 in einem neuronalen Netz. Offensichtlich existiert in dem Netz w^* , für den der Fehler geringer ist als für w_1 . Das Gradientenabstiegsverfahren wird genutzt, um diese Minima zu finden (vgl. [Son22, 52]).

(vgl. [Roj93, 251])²². Dadurch ist das Netz gegenüber *Rauschen* und *Störungen* resistent²³.

Ertel führt die Begeisterung für neuronale Netze und den Aufschwung der Neuroinformatik in den 80er Jahren wesentlich auf die "biologische Plausibilität, das gut verstandene mathematische Modell" und "die beeindruckende Simulation zur Mustererkennung" des Hopfield Modells zurück [Ert21c, 297]. *Anderson und Rosenfeld* bemerken, dass der Einfluss physikalischer Systeme auf das Hopfield Netz auch das Interesse der Physiker an neuronalen Netzen geweckt hat, und dadurch das Forschungsfeld erweitert wurde (vgl. [AR88, 458]). Dies ist darauf zurückzuführen, dass sich die Summe aller Terme in einem Hopfield-Netz wie folgt berechnet²⁴):

$$E = -1/2 \sum_{i=1}^n \sum_{j=1}^n w_{ij} x_i x_j + \sum_{i=1}^n \Theta_i x_i \quad (4.2)$$

Dies wird auch als die **Energiefunktion** des Netzes bezeichnet, wobei E entweder konstant bleibt oder kleiner wird - aber nicht größer: Ist in einem Netz eine solche Energiefunktion vorhanden, kann gezeigt werden, dass das Netz konvergiert

²² Nach dem Training ist das Netz dazu in der Lage, für einen Stimulus ein Aktualisierungsmuster einzunehmen, das diesem am ähnlichsten ist (vgl. [RN12, 882]), was an die "Cell Assemblies" von Hebb erinnert (s. a. Anhang B.2). *Lansner* verweist in [Lan09, 179] genau darauf, merkt aber gleichfalls an, dass das Modell aufgrund der Rekurrenz und Symmetrie nicht seinem biologischen Vorbild entspricht (vgl. [Lan09, 180]).

²³ Rauschen bzw. Störungen bei Stimuli in Form von Bildern ($n \times m$ Pixeln) können durch zufällig hinzugefügte oder entfernte Pixel entstehen, oder durch Rotation (axiale Verschiebung) der Daten vor dem Einspeisen in das Netz. Beispiele für verrauschte Daten in [Ert21c, 294].

²⁴ siehe [Roj93, 287]; *Rojas* weist an selber Stelle auf den Faktor $1/2$ hin: Dadurch wird die Symmetrie berücksichtigt und $w_{ij}x_i x_j$ und $w_{ji}x_j x_i$ insgesamt nur einmal berechnet wird.

(vgl. [Fau94, 139]), und zwar hin zu einem Zustand *minimaler Energie*: Ertel erklärt hierzu, dass gelernte Muster in dem Netz “Minima der Energiefunktion im Zustandsraum“ darstellen; werden zu viele Muster gelernt, “konvergiert das System gegen Minima, die keinen gelernten Muster entsprechen“. Damit ist das Modell “formal äquivalent zu einem physikalischen Modell des Magnetismus“ [Ert21c, 293]²⁵, wo ebenfalls solche *Phasenübergänge* von einem geordneten hin zu einem chaotischen System beobachtet werden können²⁶.

Ackley et al. stellen 1985 in [AHS85] die **Boltzmann-Maschine**²⁷ vor, eine Modifikation eines Hopfield-Netzes, in dem sich die Zellen *stochastisch* [AR88, 635] und globale Zustände des Netzes nach der *Boltzmann-Verteilung*²⁸ verhalten. Das Netz versucht lokale Minima zu überwinden (vgl. “Gradientenabstiegsverfahren“ in Abschnitt 4.2.1), indem ihm erlaubt wird, zu größeren E (s. Gleichung 4.2) zu springen, um so zu einem globalen Minimum zu konvergieren²⁹. Das Verfahren wird auch *simulated annealing* bezeichnet³⁰.

4.2.3 Neocognitron

Unter dem Namen *Cognitron* 1975 beschreibt *Fukushima* in [Fuk75] ein mehrschichtiges Netz mit selbst-organisierenden Eigenschaften zur Mustererkennung, in dem Zellen selektiv auf häufig präsentierte Merkmale reagieren. 1983 veröffentlichten *Fukushima et al.* eine Modifikation dieser Architektur in [FMI83]³¹ unter dem Namen **Neocognitron**³²; sein biologisches Vorbild ist das durch *Hubel und Wiesel*³³ in [HW62] beschriebene hierarchische Modell des Wahrnehmungssystems (vgl. [FMI83, 827]). In dem künstlichen neuronalen Netz haben Zellen in tiefer liegenden Schichten die Eigenschaft, selektiv komplexere Merkmale der Stimuli zu extrahieren, womit sie weniger anfällig gegenüber Rauschen in den Eingabedaten sind. In [Fuk80] war der Trainingsprozess durch wiederholte Einspeisung von Mustern ohne weitere Information gegeben (vgl. [Fuk80, 197])³⁴: Die Erweiterung des

²⁵ vgl. auch [AR98, 417] sowie [Hop82, 2556 f.]

²⁶ dem so genannten *Spinglass* (auch: *spin glass*); vgl. [BY86, 900] mit Bezug auf Hopfield-Netze

²⁷ *Hinton und Sejnowski* beschreiben das Verhalten bereits 1983 in [HS83]

²⁸ die *Boltzmann-Verteilung* gibt die Wahrscheinlichkeit an, ein geg. physik. System in einem bestimmten Zustand anzutreffen (<https://de.wikipedia.org/wiki/Boltzmann-Statistik>, abgerufen 05.09.2023)

²⁹ vgl. [AHS85, 151] sowie [Köh90, 107]

³⁰ vgl. [Ert21c, 297]. *Rojas* schreibt dazu “Gelingt es, die Temperatur nach dem richtigen Plan auf null zu verringern, wird sich das Netz mit großer Wahrscheinlichkeit in einem globalen Minimum stabilisieren.“ [Roj93, 322]. Der Begriff *Annealing* ist hier aus der Werkstoffkunde entlehnt: Eine entsprechende Wärmebehandlung von Materialien unterstützt die Erzeugung gewünschter Werkstoffeigenschaften wie Festigkeit, z.B. bei der Verarbeitung von Glas (s. [Von11, 3]) oder Schweißverbindungen von Metallen (s. [Fay+19])

³¹ mit Bezug auf [Fuk80] (“Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position“).

³² Video mit Demonstration des Netzes unter <https://www.youtube.com/watch?v=Qil4kvmv2Sw>, abgerufen 04.09.2023

³³ s. Anhang B.4

³⁴ “learning without a teacher“, also unüberwachtes Lernen (unsupervised learning).

Modells beinhaltet nun die Verstärkung der modifizierten Synapsen durch überwachtes Lernen³⁵, um bessere Resultate bei handgeschriebenen Zeichen zu erzielen (vgl. [FMI83, 829]). *Anderson und Rosenfeld* attestieren dem Netz von *Fukushima et al.* Aspekte, die bei der Entwicklung neuronaler Netze eine wesentliche Rolle spielen werden (siehe [AR88, 524]): Tatsächlich inspiriert das Neocognitron die Convolutional Neural Networks (CNN) (vgl. [LBH15, 439]), deren Architektur im Jahr 1989³⁶ Grundlage der Arbeit [LeC89] von *LeCun* ist.

4.2.4 Convolutional Neural Networks

Yann LeCun veröffentlicht 1989 seine Arbeit [LeC89], in der er verschiedene Netzwerkarchitekturen auf Generalisierungsfähigkeit (s. Abschnitt 3.2.4) und Performance untersucht. Er kommt zu dem Schluss, dass eine Reduzierung der *freien Parameter*³⁷ in einem mehrschichtigen Netz zu einer hohen Generalisierungsfähigkeit führt: Bessere Ergebnisse im Vergleich zu ein- bzw. zweischichtigen Netzen können erzielt werden, indem sog. *feature maps* genutzt werden, die in den Schichten für die Merkmalsextraktion der Eingabedaten (hier: zweidimensionale Bilder) verantwortlich sind. Zusätzliche Information wie die Lage der Merkmale in den Eingabedaten werden näherungsweise gespeichert, was zu einer Reduzierung der Größe der *feature maps* im Vergleich zu der Größe der Eingabedaten führt, und damit auch zu einer Reduzierung der Gewichte. Darüber hinaus können mehrere *feature maps* die gleichen Merkmale an unterschiedlichen Orten (*shift invariance*) in den Eingabedaten extrahieren, wodurch die Gewichte unter diesen *feature maps* geteilt werden können (vgl. [LeC89, 151 f.]). In [LeC+89] stellen *LeCun et al.* diese Architektur als **Convolutional Network** *LeNet-1*³⁸ vor (vgl. [LeC+98, 13]), das äußerst erfolgreich mittels Backpropagation handgeschriebene Postleitzahlen erkennt³⁹. Das Netz performt mit 30 Klassifizierungen pro Sekunde, lediglich die Normalisierung der Eingabedaten stellt einen Flaschenhals bei dem Prozess dar. Wird dieser berücksichtigt, werden 10-12 Ergebnisse pro Sekunde erzielt (vgl. [LeC+89, 549]).

³⁵ “We use a learning-with-a-teacher process for the reinforcement of the modifiable synapses in the new large-scale model, instead of the learning-without-a-teacher process applied to the previous model.“ [FMI83, 827]; *Anderson und Rosenfeld* bewerten dies als “some handcrafted fine tuning“ [AR88, 524 f.]

³⁶ ein Jahr nach Erscheinen von [AR88].

³⁷ *freie Parameter* sind die Parameter des Netzes, die durch den Lernvorgang festgelegt werden, also bspw. die Gewichte, Schwellenwerte oder auch die Lernrate, wenn diese durch Verwendung bestimmter Optimierungsalgorithmen adaptiv ist, wie bspw. **Adam** (vgl. [GBC18, 346] mit Verweis auf [KB17])

³⁸ siehe Anhang A.9

³⁹ Wobei die Ziffern zu Matrizen der Dimension 16×16 transformiert und als Eingabedaten in das Netz eingespeist werden. Die Transformation behält das Seitenverhältnis der Zeichen bei. Die Werte der Matrix an den Positionen a_{ij} entsprechen Grauwerten mit einem normalisierten Wertebereich von $[-1, 1]$ (vgl. [LeC+89, 542]).

Eine mathematische Basis von CNNs ist die **Faltungsoperation**⁴⁰: Hierbei wird auf Eingabedaten ein **Kernel** (die *Faltungsmatrix*)⁴¹ angewendet, wobei das Ergebnis der Faltungsoperation die **feature map** (*Merkmalskarte*) ist (vgl. [GBC18, 370]). *Goodfellow et al.* erläutern in [GBC18, 374 ff.] die Optimierungen, die durch den Einsatz von Faltung hervorgehen: Durch den Einsatz von Kernen, die nur aus einem Bruchteil der Größe der Eingabedaten bestehen⁴², können “aussagekräftige Merkmale“ aufgespürt werden, was dazu führt, dass weniger Parameter gespeichert werden müssen und gleichzeitig die “statistische Effizienz“ des Netzes erhöht wird (**sparse interaction** bzw. **sparse weights**). Durch das bereits erwähnte **parameter sharing** (auch: **tied weights**) kann die Effizienz des Netzes erhöht und Speicherplatzverbrauch verringert werden, und Schichten weisen eine **Äquivarianz** gegenüber Verschiebung auf⁴³.

CNNs nutzen neben den Convolution Schichten auch Pooling Schichten (vgl. [Ert21c, 325]), die die Ausgaben des Netzes durch eine “statistische Größe der nahegelegenen Ausgaben“ ersetzt [GBC18, 379], was die Zahl der Pixel auf den feature maps reduziert⁴⁴, außerdem wird als Aktivierungsfunktion meistens die nichtlineare ReLU (Rectified Linear Unit, s. Gleichung 4.3 und Abbildung 4.5) verwendet, die die Konvergenz der Netze verbessert (vgl. [Ert21c, 327]).

ReLU:

$$f(x) = \max(0, x) \quad (4.3)$$

⁴⁰ “convolution“ (engl.): Faltung

⁴¹ “das biologische Analogon zu den Filtern [Kernen] sind die bei den Sinnesorganen vorkommenden rezeptiven Felder.“ [Ert21c, 326]. Vgl. hierzu auch [LBH15, 439].

⁴² *Goodfellow et al.* beziffern für Eingaben von 1.000.000 Pixeln Kernel mit einer Größe von einigen hundert Pixeln als ausreichend (siehe [GBC18, 374]).

⁴³ bereits erwähnt wurde die *shift invariance*: Wird bspw. ein Bild X in das Netz f eingespeist, das in einer Ausgabe $f(X) = Y$ resultiert, dann bedeutet *shift invariance*, dass eine Verschiebung der Eingabedaten $g(X) = X + \Gamma$ in dem gleichen Y resultiert: $f(X) = f(g(X))$. *Äquivarianz* (engl. “equivariance“) bedeutet, dass eine Änderung der Eingabedaten auch zu einer “gleichartigen Änderung der Ausgabe“ führt: $f(g(x)) = g(f(x))$ (vgl. [GBC18, 377]).

⁴⁴ *LeCun et al.* weisen darauf hin, dass die Pooling Schicht auch die Aufgabe hat, semantisch ähnliche Merkmale zusammenzufassen: “the role of the pooling layer is to merge semantically similar features into one.“ [LBH15, 439]

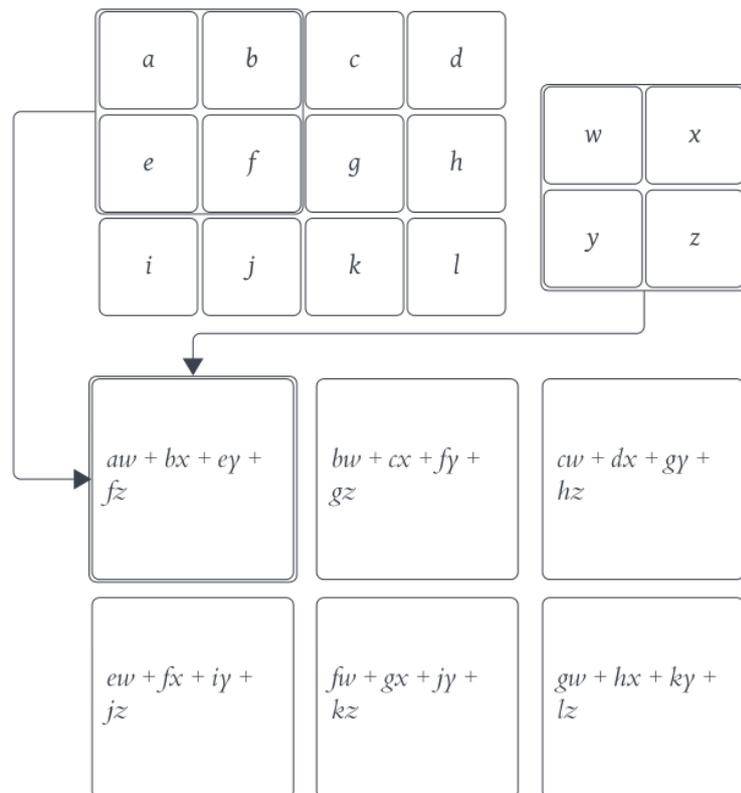


Abb. 4.4. Beispiel einer Faltungsoperation. (Quelle: in Anlehnung an [GBC18, 372, Abbildung 9.1])

In dem Beispiel wird zur Faltung ein 2×2 -Kernel mit einer 3×4 -Eingabe verwendet. Das Ergebnis ist eine 2×3 -feature map.

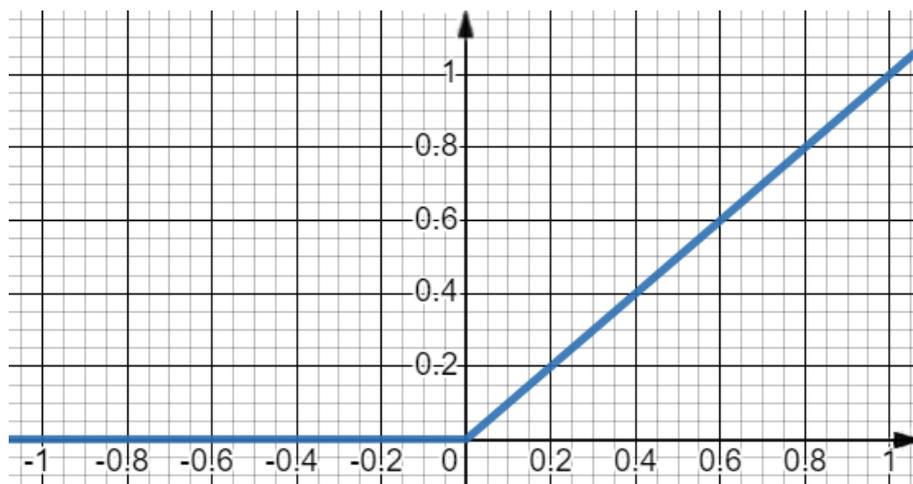


Abb. 4.5. Plot der ReLU. (Quelle: Eigene Darstellung)

Neuronale Netze im Gesundheitswesen

Pfannstiel stellt in [Pfa22, 35, Abb. 1.14] die Rolle der Künstlichen Intelligenz im Gesundheitswesen als Schichtenmodell dar: Künstliche Intelligenz wird umgeben von vier Schichten, von denen jede eine Menge repräsentativer Leistungen enthält. Diese lassen sich grob einteilen in¹

1. Versorgung des Patienten
2. Herstellung / Zulieferung durch Pharmazie bzw. Medizintechnologie
3. Organisation des Gesundheitswesens
4. Selbstversorgung des Patienten

In jeder dieser Schicht übernimmt die KI wichtige Rollen, z.B. in der Diagnostik (1.), der Wirkstoffentwicklung (2.), der Abrechnung von Leistungen (3.) oder als Software in embedded Systems wie Wearables (4.). Zum Beispiel werden die in Abschnitt 4.2.4 vorgestellten CNNs aufgrund ihrer Erfolge bei der Objektklassifikation² in der Diagnostik eingesetzt³: Je nach Verfahren zeigt sich, dass die Netze mindestens genauso gute Diagnosen erstellen wie erfahrene Ärzte, und bessere als wenig erfahrene Ärzte (vgl. [She+19, 7]). Doch auch in anderen Bereichen des Gesundheitswesens finden neuronale Netze praktische Anwendungen. Im Folgenden werden einige vorgestellt.

¹ Reihenfolge analog zu dem Modell in [Pfa22], von Innen nach Aussen

² vgl. [Ert21c, 330]. *Goodfellow et al.* stellen etwas allgemeiner fest, dass CNNs “gewaltige Erfolge in der Praxis gebracht“ haben [GBC18, 369]. Als Beispiel sei PSPNet genannt, ein CNN zur semantischen Segmentierung von Bilddaten (“*scene parsing*“), das *Zhao et al.* in [Zha+17] vorstellen. Heutzutage betreiben Unternehmen wie Google, Facebook, Microsoft und IBM Forschung und Entwicklung im Bereich der Bildverarbeitung mit CNNs, und im Bereich Hardware entwickeln bspw. NVIDIA (<https://nvidianews.nvidia.com/news/nvidia-introduces-drive-agx-orin-advanced-software-defined-platform-for-autonomous-machines>, abgerufen 07.09.2023) und Samsung Chips zur Echtzeiterkennung der Umgebung [LBH15, 440].

³ bspw. stellen *Brinker et al.* in [Bri+18] 13 Arbeiten zur Klassifikation von Hautläsionen mit Hilfe von CNNs vor.

5.1 Automatisierung in der Gesundheitswirtschaft

Um für die in Deutschland knapp 73 Millionen gesetzlich Krankenversicherten⁴ Leistungen im Gesundheitswesen abzurechnen, werden durch die Leistungserbringer Finanzdienstleister beauftragt. So leiten bspw. Apotheken eingelöste Kassenrezepte an Apothekenrechenzentren weiter, die die Abrechnungen mit den Krankenkassen durchführen. Obwohl Rezepte heutzutage überwiegend maschinell erstellt werden, kommt es bei den Rechenzentren nach der Digitalisierung der Rezepte (bspw. durch Scannen) häufig zu manueller Nachkorrektur, wenn abrechnungsrelevante Daten nicht vollständig erfasst werden können, wie *Höfer et al.* in [HWN22, 698] feststellen. Um den zusätzlichen Arbeitsaufwand durch Nachkorrekturen zu reduzieren (aufgrund von Störfaktoren wie ausgedünnte Zeichen oder Textüberlagerungen), zeigen *Höfer et al.* ebenda ein neuronales Netz, das in Auftrag eines führenden Abrechnungs- und Finanzdienstleisters erstellt wurde. Ziel des Netzes ist die Rekonstruktion⁵ von Rezepten, die bei Plausibilitätskontrollen auffällig sind.

Ihr neuronales Netz zur Rekonstruktion von Rezepten ist ein CNN mit 7 Faltungsschichten, denen jeweils eine Poolingschicht folgt. Insgesamt werden 40.000 Datensätze für das Training genutzt, sowie 10.000 Validierungsdatsätze. Die Rezepte werden auf den für die Abrechnung relevanten Teil zugeschnitten, sodass die Eingabedaten aus 720×460 Pixeln bestehen. Sie zeigen, dass Ihr Netz bei der Rekonstruktion von Korrekturzeichen eine relative Verbesserung von 27,85% erreicht⁶ und aufgrund dieses Erfolges in den Workflow der zu dem Auftraggeber gehörenden Abrechnungszentren integriert wurde.

5.2 Pharmaforschung

Im Jahr 2016 gelingt es der von *Google DeepMind* entwickelten Software *AlphaGo*, den Südkoreaner Lee Sedol als einen der weltbesten *Go*-Spieler in diesem Spiel zu schlagen⁷. *Go* verfügt über $2,08 \times 10^{170}$ gültige Spielpositionen und gehört damit zu den komplexesten Brettspielen⁸. Kurz nach diesem beachtlichen Erfolg entscheidet *Google DeepMind*, dass das in der Software verwendete 13-lagige CNN-Netz mit seiner 19×19 Eingabe- und Ausgabematrix sowie trainiert mit über 30 Millionen Spielzügen (vgl. [Ert21a, 371]) reif sei für den Einsatz in der Wissenschaft. Im selben Jahr beginnt das Unternehmen mit der Forschung an *AlphaFold*⁹: Die Software soll helfen, das Problem der **Proteinfaltung** zu vereinfachen. Darunter

⁴ https://www.gkv-spitzenverband.de/krankenversicherung/kv_grundprinzipien/alle_gesetzlichen_krankenkassen/alle_gesetzlichen_krankenkassen.jsp, abgerufen 12.09.2023

⁵ tatsächlich liegt die Vermutung nahe, das Netz sollte eingesetzt werden, um den Hochleistungs-scanner zu ersetzen. Ziel war es aber, in einem Vorverarbeitungsschritt die Bildqualität zu erhöhen (vgl. [HWN22, 698])

⁶ 5% waren ursprünglich anvisiert (vgl. [HWN22, 711]).

⁷ <https://www.spiegel.de/netzwelt/gadgets/go-duell-software-alphago-siegt-gegen-lee-sedol-a-1081975.html>, abgerufen 08.09.2023

⁸ <https://trompt.github.io/go/legal.html>, abgerufen 08.09.2023

⁹ [Jum+21], s. a. <https://www.deepmind.com/research/highlighted-research/alphafold/timeline-of-a-breakthrough>, abgerufen 08.09.2023

versteht man die Vorhersage der Struktur eines Proteins auf Basis seiner Aminosäuresequenz (vgl. [Dil+08])¹⁰; mit 10^{300} verschiedenen Konstellationen¹¹ ist die Struktur schwer zu berechnen. Der Pharmaindustrie erlaubt das Wissen über die Struktur von Proteinen die Herstellung von Medikamenten, deren Wirkstoffe bspw. Proteinfunktionen aktivieren oder hemmen¹².

Großen Erfolg hat die Software 2020 bei dem CASP14-protein-folding-contest¹³. Mit 170.000 Proteinstrukturen trainiert und von ~ 200 GPUs unterstützt erreicht AlphaFold2 Bestwerte¹⁴, was als Durchbruch für die Medizin bewertet wird (vgl. [Cal20, 204]).

5.3 Diagnostik

Szlovits et al. argumentieren 1988 in [Szo88], dass die Abfrage von Symptomen zur Feststellung von Krankheiten¹⁵ auch von Expertensystemen (s. Abschnitt 4.1) übernommen werden kann. Es zeigt sich dort ebenfalls, dass Ärzte kritisch gegenüber computergestützten Assistenzsystemen stehen. *Teach und Shortliffe* stellen bereits 1981 fest, dass es medizinischem Fachpersonal wichtig ist, dass das System den Entscheidungsweg zur Diagnose erklären kann; dass der Computer stets die korrekten Diagnosen stellt, wird als weniger wichtig bewertet (vgl. [TS81, 551, “Table V“: “D.1“ sowie “D14“]. In dieser Studie werden auch ethische Bedenken seitens der Mediziner aufgeführt¹⁶. Dass computergestützte medizinische Assistenzsysteme dann kaum Anwendung im klinischen Umfeld fanden, führen *Lucieri et al.* in [Luc+22] auch auf solche Bedenken zurück: Erst steigende Rechenleistung und die fortschreitende Leistungsfähigkeit tiefer neuronaler Netze weckte das Interesse an KI unter den Mediziner, dank der Erfolge bildbasierter CNNs auch in der Diagnostik (vgl. [Luc+22, 728]). Dem oben erwähnten Wunsch einer **erklärbaren KI** (**xAI**¹⁷) wird seitdem nachgegangen, was sich aufgrund des *Blackbox-Charakters*¹⁸

¹⁰ postuliert 1972 von *Anfinsen* in seiner Nobelpreislaudatio [Chr73, 223]

¹¹ <https://web.archive.org/web/20110523080407/http://www-miller.ch.cam.ac.uk/levinthal/1evinthal.html>, abgerufen 08.09.2023

¹² Bei dem strukturbasierten Wirkstoffdesign hilft die Kenntnis über die Struktur des Zielproteins bei der Entwicklung von Medikamenten, deren Moleküle zur Wirkstoffentfaltung an die Proteine andocken (vgl. [SKM10, 29 ff.]).

¹³ CASP (Critical Assessment of Techniques for Protein Structure Prediction), ein seit 1994 alle zwei Jahre stattfindender Wettbewerb zur Vorhersage von Proteinstrukturen (<https://predictioncenter.org/>, abgerufen 08.09.2023)

¹⁴ mit einem Score von 92.4 GDT. In den Jahren vor AlphaFold lag der Score bei ca. 30-40 GDT. GDT bedeutet *global distance test* mit einem Wertebereich zwischen 1 und 100 und gibt die Ähnlichkeit einer Proteinstruktur mit der Struktur eines Vergleichsmodells an (https://proteopedia.org/wiki/index.php/Calculating_GDT_TS, abgerufen 09.09.2023)

¹⁵ bei gleichzeitigem Ausschluss anderer Krankheiten

¹⁶ In dem Papier findet sich auch die Befürchtung unter den Teilnehmern an der Umfrage, dass computergestützte Systeme den Arzt ersetzen könnte (vgl. [TS81, 543]).

¹⁷ Abkürzung für (engl.) *explainable AI*

¹⁸ <https://spectrum.ieee.org/ai-failures>, abgerufen 09.09.2023

insbesondere bei Architekturen mit Millionen von Parametern¹⁹ in neuronalen Netzen aber als schwierig erweist²⁰.

Trotz allem ist nicht von der Hand zu weisen, dass neuronale Netze in der Qualität der Diagnosen gleichauf sind mit denen von medizinischem Fachpersonal (vgl. [She+19, 1]). *Amato et al.* listen in [Ama+13] unter anderem Studienergebnisse bzgl. des Einsatzes neuronaler Netze bei der Diagnostik von kardiovaskulären Krankheiten, Tumor-Erkrankungen sowie Diabetes auf, und kommen bei der Auswertung zu dem Ergebnis, dass die neuronalen Netze für einer Vielzahl verschiedener Symptome korrekte Diagnosen erstellen: Zum Beispiel werden Audioaufnahmen der Pumpfähigkeit des Herzens zur Klassifizierung von Herzklappenfehler genutzt, bei dem das eingesetzte Netz - ein MLP mit 3-Schichten unter Verwendung von Backpropagation - eine Erfolgsrate von knapp 95% vorweisen kann²¹.

In [Est+17] trainieren *Estava et al.* ein CNN basierend auf *GoogleNet Inception v3*²² mit über 1.000.000 Bildern aus 1000 verschiedenen Kategorien, um danach über *transfer learning*²³ mit 129.450 gelabelten Bilddateien aus 2032 verschiedenen Krankheiten ein Netz zur Diagnose von Hautkrebs zu erstellen. Sie zeigen, dass das Netz bei der Korrektheit der gestellten Diagnosen genauere Diagnosen erstellt als ein Mediziner²⁴.

Irving et al. stellen in [Irv+19] "CheXpert"²⁵ vor, einen öffentlichen Datensatz mit 224.316 Aufnahmen des Thorax von 65.240 verschiedenen Patienten zur Klassifizierung von 14 Befunden (u.a. Lungenentzündung, Fraktur, Pleuraerguss). Für das Training nutzen sie *DenseNet121*, ein CNN, in dem jede Schicht mit jeder anderen verbunden ist (vgl. [HLW16])²⁶. Die Inputdaten bestehen aus 320×320 Pixeln. Bei den Tests zeigt sich, dass ihr Netz bei der Diagnose von Kardiomegalie, Ödemen sowie Pleuraergüssen besser abschneidet als die Radiologen, die für die Studie zum Vergleich eingesetzt wurden²⁷.

¹⁹ **AlexNet**: 60.000.000 Parameter [KSH12, 1]

²⁰ Einige erfolgreiche Ansätze jüngster Zeit fassen *Lucieri et al.* in [Luc+22, 733 ff] zusammen. *Steinwender* geht in [Ste22, 765] auf die Zertifizierungspflicht von Medizinprodukten in der EU ein, und das mit der damit verbundenen Datenschutzgrundverordnung auch eine Interpretierbarkeit bzw. Erklärbarkeit für die KI-Systeme einhergeht

²¹ Bezug auf [Uğü12, 71]

²² Im Vergleich zu **AlexNet** nur 5.000.000 Parameter, aber erreicht bei der Bildklassifizierung bessere Werte als andere Modelle (vgl. [Sze+15]). Eine interessanten Diskussionsgrundlage zu den Ausführungen in [LeC89] (s. Abschnitt 4.2.4).

²³ beim *transfer learning* werden semantisch gleiche Ausgaben für verschiedenklassige Eingaben genutzt (vgl. [GBC18, 602 f.]).

²⁴ "The deep learning CNN outperforms the average of the dermatologists at skin cancer classification using photographic and dermoscopic images" [Est+17, 3, Figure 2]

²⁵ <https://stanfordmlgroup.github.io/competitions/chexpert/>, abgerufen 10.09.2023

²⁶ DenseNet: Dense Convolutional Network. In klassischen CNNs mit L Schichten existieren L direkte Verbindungen, also eine Verbindung für eine obere und die darunterliegende Schicht. Durch die Architektur von *Densenet* existieren in dem Netz $L(L+1)/2$ direkte Verbindungen.

²⁷ siehe auch <https://intermountainhealthcare.org/news/2019/09/ai-system-accurately-detects-key-findings-in-chest-x-rays-of-pneumonia-patients-within-10-seconds/> (abgerufen 10.09.2023) fasst die Ergebnisse bei der Diagnose von Lungenentzündung zusammen

5.4 Therapie und Prognose

Die *HL7*-Organisation²⁸ wurde 1987 mit dem Ziel gegründet, einen Kommunikationsstandard für den elektronischen Datenaustausch im Gesundheitswesen zu etablieren. Teil des Standards ist die *FHIR*-Spezifikation²⁹, die von *Rajkomar et al.* in [Raj+18] für ein neuronales Netz genutzt wird, das für einen mit diesen Daten verknüpften Patienten bei der Hospitalisierung Vorhersagen über folgende Punkte erstellt:

- Mortalität während des Krankenhausaufenthaltes (“inpatient mortality“)
- ungeplante stationäre Wiederaufnahme (“unexpected readmissions within 30 days“)
- Verlängerung des Aufenthaltes (“long length of stay“)
- Befund bei Entlassung (“discharge diagnoses“)

Validiert wurde das Netz über insg. 46.864.534.945 Datenpunkte: Die von verschiedenen Krankenhäusern zur Verfügung gestellten 216.221 Patientenakten wurden hierzu aufgeteilt in 194.470 Trainingsdaten und 21.751 Testdaten, allesamt im *FHIR*-Format. Die Autoren des Papiers konnten zeigen, dass ihr Netz die traditionellen klinischen Modelle in den ersten 3 genannten Punkten übertrifft.

Eine ähnliche Arbeit stellen *Choi et al.* in [Cho+16] vor, in der sie untersuchen, ob elektronische Patientendaten genutzt werden können, um mit Hilfe rekurrenter neuronaler Netze Vorhersagen zu Diagnose, Medikation und Wiedervorstellung beim Arzt zu treffen. Sie zeigen, dass ihr Netz als medizinisches Assistenzsystem in den aufgeführten Punkten geeignet ist, und darüber hinaus mittels *transfer learning* in Kliniken eingesetzt werden kann, bei denen die Datenmenge nicht für ein Training des Netzes ausreicht.

Prognosemodelle werden auch auf anderer Ebene genutzt, bspw. bei der Evaluierung der korrekten Behandlungsmethode. In [Li+21] stellen *Li et al.* *G-Net* vor, ein Deep Learning Framework für “counterfactual prediction“ als Assistenzsystem für Ärzte. Es soll dabei helfen, unter Berücksichtigung von zur Verfügung stehenden Daten wie Therapie- und Krankheitsverlauf der Patienten die richtige Behandlungsstrategie auszuwählen. Unter “counterfactual prediction“ sind hier alternative Szenarien gemeint, also das Ergebnis einer zunächst in der Theorie eingeschlagenen Therapie, für die von der Software das Behandlungsergebnis ermittelt wird³⁰.

Als Beispiel führen *Li et al.* die Flüssigkeitszufuhr zur Stabilisierung von Intensivpatienten an, die einen septischen Schock erlitten haben. Anwendungszeitpunkt, -art und -menge bestimmen den Genesungsverlauf und können unerwünschte Nebenwirkungen verhindern, allerdings auch solche begünstigen, die bis hin zum Tode führen³¹. Die in *G-Net* verwendeten Netze sollen dabei helfen, in Verbindung

²⁸ “Health Level Seven“ <https://hl7.org>, abgerufen 11.09.2023

²⁹ “Fast Healthcare Interoperability Resources“, <https://hl7.org/fhir/>, abgerufen 11.09.2023

³⁰ “This is particularly important in clinical settings, where physicians may have to choose between multiple treatment strategies for their patients but are unable to test all of them before making a decision“ [Li+21, 282]

³¹ vgl. [Sha+20, 5 f.]

von Parametern wie Laborwerte des Patienten sowie Zeitpunkt und Menge der Verabreichung eine Aussage über den Behandlungseffekt zu treffen³², um so die bestmögliche Behandlungsmethode auszuwählen.

³² unter <https://news.mit.edu/2022/deep-learning-technique-predicts-clinical-treatment-outcomes-0224> (abgerufen 10.09.2023) ist eine ausführlichere Pressemitteilung mit Verweis auf das zitierte Papier zu finden

Zusammenfassung und Bemerkung

In Kapitel 2 haben wir gesehen, dass ein biologisches neuronales Netz ein Kommunikationsverbund ist, in dem hochkomplexe biochemische Prozesse ineinandergreifen. Diese Prozesse ermöglichen eine Informationsverarbeitung: Eine einzelne Nervenzelle ist in der Lage, Signale zu empfangen, zu verrechnen und ein Signal auszugeben. Löst ein Aktionspotenzial aus, diffundieren Neurotransmitter in den synaptischen Spalt, die Folgezelle verrechnet die daraus entstehenden elektrischen Signale und so setzt sich die Kette fort, oft bis tausende von Neuronen ein Signal an das Ende der Kommunikationskette weitergeleitet haben. Ob ein einzelnes Vesikel ein **mEPSC** verursacht, mag dabei auf analytischer Ebene interessant sein. Die Frage, welches der oft über tausend Vesikel, die in einem Neuronenverband an einer Signalweiterleitung beteiligt sind, ursächlich für eine muskuläre Reaktion ist, ist sicherlich schwieriger zu beantworten.

Tiefe künstliche neuronale Netz mit ihren Millionen von Parametern (vgl. [Sze+15]) erweisen sich im Hinblick auf eine Analyse ähnlich undurchschaubar¹.

Die Modellierung solcher Netze folgt dabei in Teilen dem biologischen Vorbild, und es mag zunächst überraschen, dass einige solcher Verschaltungsmuster - wie die **Divergenz** oder die **rekurrente** oder **laterale Hemmung** - zunächst an Logikgatter erinnern (vgl. [Eil19, 58 f.]). Es liegt nahe, dass die Wissenschaft versucht ist, solche naturgegebenen verrechnenden Einheiten und logischen Verschaltungen zu verstehen und in mathematische Formeln zu gießen, um ein Modell zu erstellen, das menschliche Intelligenz erklärt und nachstellt, und dann als Abstraktionen auf Maschinen übertragen wird: “if you really understand something, you can usually make a machine do it“ [AR88, xiii].

Einer dieser Versuche wurde in Abschnitt 3.1 mit dem MCP-Modell vorgestellt, ein **empirisches Modell**², das auf Analyse und einfacher Schwellenwertlogik basiert (vgl. [AR88, 16]).

¹ dieser Blackbox-Charakter wird problematisch, wenn bspw. von Assistenzsystemen in der Medizin Entscheidungswege aufgezeigt werden sollen, die zu einer bestimmten Diagnose oder Therapieempfehlung führen (siehe Abschnitt 5.3).

² auch “caricature model“; beides in [HI97, 4]

In dem dort zitierten Papier weisen McCulloch und Pitts darauf hin, dass das Alles-Oder-Nichts-Prinzip (vgl. Abschnitt 2.3.1) als Vorbild für ihre Abstraktion durch Aussagenlogik dient. Das von ihnen erstellte Modell ist dabei eine starke Vereinfachung des biologischen Vorbilds und weist dadurch eine äußerst nützliche Eigenschaft auf:

This ensures its status as a *model*, and not a *copy*, of a real neuron, and makes it possible to implement on a digital computer. ([BJ90, 43; Hervorhebung i. O.]).

Dabei wurde ein anderer neurobiologischer Prozess, der heutzutage bei maschinellen Lernverfahren Standard ist, nicht von ihnen berücksichtigt: Die dynamische Anpassung des Netzes³.

McCulloch und Pitts sind sich all dessen durchaus bewusst⁴. Entsprechend ist ihr Modell statisch und muss zur Bewältigung von Aufgaben zunächst durch Änderung der Netztopologie bzw. Anpassung einzelner Schwellenwerte “konstruiert” werden (vgl. [Fau94, 28] sowie [Roj93, 51]), selbständig zu lernen vermag es nicht.

Die McCulloch-Pitts-Zelle ist von eher geringer Bedeutung für die Neurowissenschaft gewesen⁵. Umso größer war ihr Einfluss auf die Computerwissenschaften⁶: Logische Schaltungen, die Intelligenz und Kognition erklären sollen, brachten den Gedanken an eine “intelligente Maschine” hervor (vgl. [Pic04, 204]). So war diese Formalisierung ein wichtiges Schlüsselement für die von-Neumann-Rechnerarchitektur⁷ sowie Wieners Kybernetik⁸, und zusammen mit den nachfolgenden Arbeiten von *Hebb*⁹ und *Rosenblatt* ebnete es den Weg für die Forschung an künstlicher Intelligenz (vgl. [Arb00, 1]).

Erwähnenswert ist auch die Bemerkung von McCulloch und Pitts bzgl. “Berechenbarkeit“:

This is of interest as affording a psychological justification of the Turing definition of computability and its equivalents, Church’s A-definability and Kleene’s primitive recursiveness: if any number can be computed by an

³ “Lernen“ geschieht auf physiologischer Ebene durch die Modulierung von Synapsen (vgl. [HS19c, 115]); McCulloch und Pitts setzen in ihrem Kalkül voraus, dass sich das Netz nicht verändert (siehe Abschnitt 3.1.1).

⁴ “McCulloch and Pitts acknowledged in their paper that their definition of a neuron was idealized, and that they made physical assumptions that were ‘most convenient for the calculus’“ [Abr02, 21]. Siehe auch [MP43, 101]: “we regard facilitation and extinction as dependent upon continuous changes in threshold related to electrical and chemical variables“, sowie “He [McCulloch] never claimed that the 1943 model exhausted the richness of individual neurons“ [Arb00, 11]

⁵ “The immense theoretical influence of this paper was not among neuroscientists but among computer scientists.“ [AR88, 17]; *Rojas* kommt zu einem ähnlichen Schluss: “Die Art der Schaltungen, die mit diesen Zellen gebaut werden, ist aber biologisch gesehen nicht so relevant.“ [Roj93, 51]

⁶ Die Arbeit war ebenso wichtig für die Entwicklung des “Konnektionismus“(vgl. [Arb00, 11]), einer Forschungsrichtung der künstlichen Intelligenz, in der Modelle (neuronale Netze) untersucht werden, mit deren Hilfe sich intelligente und kognitive Handlungen auf Maschinen übertragen lassen (vgl. [Dor91, v])

⁷ siehe [von93]

⁸ siehe Anhang B.8

⁹ siehe [Heb49]

organism, it is computable by these definitions, and conversely. ([MP43, 113])

Sie nehmen hiermit Bezug auf die von *Alan Turing* (1912 - 1954) bereits 1936 veröffentlichte Arbeit [Tur37], in der Turing die Beschreibung der Operationen seines Computers einleitet mit:

Let us imagine the operations performed by the computer to be split up into 'simple operations' which are so elementary that it is not easy to imagine them further divided. ([Tur37, 250])

Parallelen zu der Architektur der von Turing aufgestellten Modelle und dem von McCulloch entworfenen "psychon" als kleinste Einheit psychischer Aktivitäten sind erkennbar. *Piccinini* stellt dazu fest, dass die Aussage, MCP-Netze können Berechnungen anstellen, einen ersten wichtigen Bezug zwischen mathematischer Theorie und Neurowissenschaft herstellte (vgl. [Pic04, 197]).

McCulloch schreibt dazu später

Pitts and I showed that brains were Turing machines, and that any Turing machine could be made out of neurons. ([McC16, 164])

und stellt fest, dass das Nervensystem eine ausgezeichnete logische Maschine repräsentiert (vgl. [McC16, 80]).

Konsequenterweise nutzt *Minsky* in [Min67, 32 ff.] die MCP-Zelle als *endlichen Automaten*^{10,11} mit zwei Zuständen, aus denen andere endliche Automaten gebaut werden, und nennt diese, im Jahr 1967 schon geläufig: **Neuronale Netze** [Min67, 33].

Minsky analysiert ebenfalls das im Abschnitt 3.2 vorgestellte Perzeptron.

In [MP88] untersucht er zusammen mit *Papert* u.a. das Verhalten des Perzeptrons im Fall nicht-separierbarer Daten¹² sowie das Problem bzgl. *recognition of connectedness*¹³, bei dem es um die Erkennung zusammenhängender geometrischer Figuren geht.

¹⁰ Ein endlicher Automat ist eine informationsverarbeitende Maschine. Basierend auf Eingaben lassen sich mit einer solchen Maschine Zustände, Zustandsübergänge und Ausgaben modellieren. Bei einem *endlichen Automaten* sind die genannten Eingaben, Zustände, und Ausgaben endlich (vgl. [SSH95, 26 ff.]).

¹¹ dass MCP-Netze endliche Automaten sind: Auch in [Cow90, 76] sowie [Roj93, 47, "Satz 2.4"]. *Arbib* weist darauf hin, dass die Arbeit von McCulloch und Pitts grundlegend für Automatentheorie gewesen ist (vgl. [Arb00, 8]).

¹² in [MP88, 181 ff.]

¹³ s. [MP88, 12, "Theorem 0.8"] sowie [MP88, 249 f.]. In dem Epilog, der 19 Jahre nach der Erstveröffentlichung geschrieben wird, machen sie deutlich, dass sie nicht nachweisen wollten, dass ein Perzeptron nicht in der Lage zu der Erkennung zusammenhängender geometrischer Figuren sei, sondern dass die Komplexität von verschiedenen Aufgaben Anforderungen an ein Perzeptron stellt, die zu der damaligen Zeit (Ende 1950er / Anfang 1960er) schwierig umzusetzen gewesen sind (vgl. [MP88, 250])

Ihr Anliegen mit dem Buch ist es, die (mathematischen) Grenzen des Rosenblatt-Perzeptrons auf ein formales Gerüst zu stellen (vgl. [MP88, 249]). Allerdings verstärken ihre Ausführungen die zu diesem Zeitpunkt ohnehin schon skeptische Haltung¹⁴ gegenüber der Fähigkeiten des Perzeptrons¹⁵ und künstlicher neuronaler Netze allgemein, was *Anderson und Rosenfeld* rückblickend auch auf Aussagen in der Einleitung des Buches zurückführen wie¹⁶

hundreds of projects and experiments [bzgl. des Perzeptrons] were generally disappointing, and the explanations inconclusive. The machines usually work quite well on very simple problems but deteriorate very rapidly as the tasks assigned to them get harder. ([MP88, 19])

Diesen “Pessimismus“ betrachten *Minsky und Papert* im Jahr 1988 in Retrospektive (vgl. [MP88, xiii]). Sie sind sich der Behauptung bewusst, ihr Buch hätte mit dem Aufzeigen der Grenzen des Rosenblatt-Modells den Forscherdrang an maschinellem Lernen gebremst:

One popular version is that the publication of our book so discouraged research on learning in network machines that a promising line of research was interrupted. ([MP88, xii])

Anderson und Rosenfeld erwähnen “the influence of Marvin Minsky and Seymour Papert on the loss of interest in neural networks during the 1970s“ [AR98, X].

Russell und Norvig fassen zusammen, dass *Minsky und Papert* in ihrem Buch bewiesen haben, dass ein Perzeptron alles lernen kann, was es auch darstellen kann, aber es könnte halt nur sehr wenig darstellen (vgl. [RN12, 45])¹⁷. Auf gleicher Seite verweisen sie auf die zunehmende Komplexität der zu berechnenden Modelle, die nicht alleine durch schnellere und bessere Hardware kompensiert werden konnte.

Der Lighthill Report¹⁸ wird 1973 die britischen Regierung dazu bewegen, das Budget für die Forschung an KI zu kürzen (vgl. [RN12, 45]): Neuronale Netze werden als Grundlage künstlicher Intelligenz zunächst verworfen¹⁹, und Forschungsarbeiten an ihnen geht bis zum Anfang der 1980er Jahre zurück. Diese Periode ist gemeinhin als “KI-Winter“ (siehe Anhang A.8) bekannt.

Die in Kapitel 4 vorgestellten Architekturen und Algorithmen sorgten dann zusammen mit dem technologischen Fortschritt für ein erneutes Aufleben der Forschung an neuronalen Netzen und Künstlicher Intelligenz in den 1980er Jahren.

¹⁴ Here was a machine that could do pattern recognition in a humanlike way; it could recognize all kinds of things. Almost everyone at MIT was very skeptical“ [AR98, 99]

¹⁵ auch wegen schlechter Skalierbarkeit des Modells in der Praxis (vgl. [AR88, 159])

¹⁶ Auch eine generelle Irritation und Enttäuschung über das Perzeptron-Modell in Hinblick auf mehrschichtige Modelle meinen *Anderson und Rosenfeld* in “we [Minsky und Papert] consider it to be an important research problem to elucidate (or reject) our intuitive judgement that the extension is sterile.“ [MP88, 232] zu erkennen (vgl. [AR88, 159]).

¹⁷ vgl. hierzu [MP88, xiii]

¹⁸ “Workers entered the field around 1950, and even around 1960, with high hopes that are very far from having been realised in 1972. In no part of the field have the discoveries made so far produced the major impact that was then promised.“ in https://www.chilton-computing.org.uk/inf/literature/reports/lighthill_report/p001.htm, “3 Past disappointments“, abgerufen 28.08.2023

¹⁹ vgl. [Ola96, 641]

Einige der Ergebnisse dieser Anstrengungen wurden in Kapitel 5 in Auszügen vorgestellt. Dass dabei die Algorithmen nicht immer der Funktionsweise des natürlichen Vorbildes entsprechen, haben wir bei dem Backpropagation-Verfahren gesehen (siehe Abschnitt 4.2.1).

Bei allem Fortschritt, der durch Forschung und Wissenschaft in den vergangenen Jahren erreicht wurde, ist bei den hier vorgestellten Ergebnissen aber vor allem eins deutlich: Die hohe Leistungsfähigkeit einiger dieser Netze kann nur durch Lernen erreicht werden (vgl. [AHR19, 40]), wozu umfangreiche, qualitativ hochwertige Daten benötigt werden.

Nguyen und Patrick stellen in [NP16] fest, dass klinische Daten aufgrund verschiedener Faktoren wie Heterogenität (maschinell / handschriftlich erstellte Daten) und weiterem Rauschen, wie unbekanntem Abkürzungen oder Rechtschreibfehlern, als Trainingsdaten für maschinelles Lernen eine besondere Herausforderung darstellen.

Normierte Daten können deshalb helfen, die Auswertung, Anwendung aber auch den Austausch derselben zu erleichtern, wie das Beispiel aus Abschnitt 5.4 zeigt (vgl. auch [AHR19, 42]).

Wir dürfen schließen, dass Künstliche Intelligenz in Form neuronaler Netze nicht nur durch komplexe mathematische Modelle und hohe Rechenleistung ermöglicht wird: Sie benötigt zu der erfolgreichen Bewältigung ihrer Aufgaben auch eine gesunde Datenbasis für ihre Lernverfahren²⁰.

²⁰ *Dash et al* fassen in [Das+19] einige wesentliche Punkte zum Thema Big Data im Gesundheitswesen zusammen, ausserdem *Rüping und Sander* in [RS19], die dort feststellen: “Bei der datengetriebenen Entwicklung von Systemen ist zu beachten, dass die Qualität des Ergebnisses zentral von der Qualität der Eingabedaten abhängt.“ [RS19, 19]

Anhänge

A

Begriffe und Ergänzungen

A.1 Aktionspotenzial

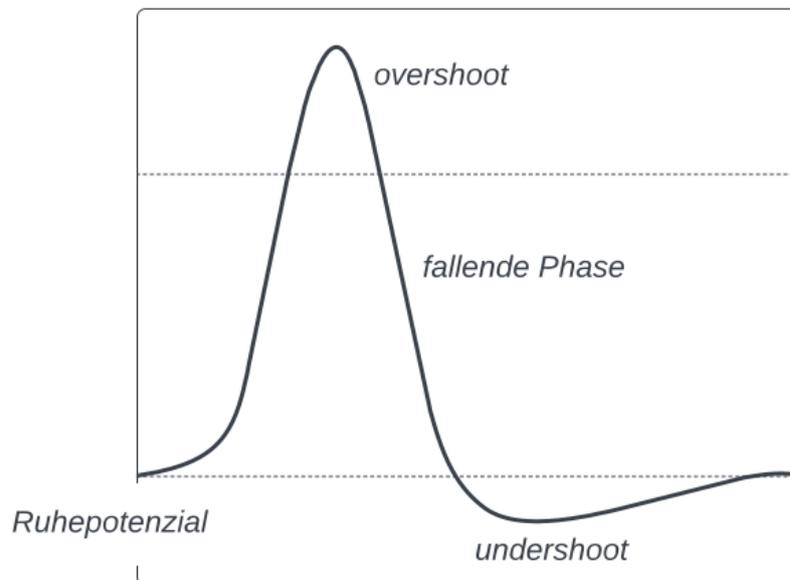


Abb. A.1. Phasen des Aktionspotenzials. (Quelle: in Anlehnung an [BCP18, 88, Abb. 4.1b])

Phasen des Aktionspotenzials

Die Entstehung eines Aktionspotenzials kann entsprechend der zeitlichen Reihenfolge in folgende Phasen eingeteilt werden:

1. Aufstrich

Na^+ gelangt in die Zelle, g_{Na} wird erhöht, mehr Na^+ strömt ein (positiven Rückkoppelung). Es kommt zu einer Depolarisation, V_m wächst exponentiell (vgl. [SD07, 46]).

2. Overshoot

V_m wird positiv ($> 0mV$) und nähert sich insgesamt $E_{Na} \sim 60mV$ an¹ (vgl. [BCP18, 105]).

3. Repolarisationsphase

Auch: Fallende Phase (vgl. [BCP18, 105]). Natriumkanäle werden **inaktiviert**. Spannungsabhängige Kaliumkanäle öffnen sich ca. $1ms$ nach Depolarisation², K^+ -Ionen strömen aufgrund der elektrochemischen Triebkraft in den EZF, das Membranpotenzial wird wieder negativ³.

4. Undershoot⁴

Es kommt zu einem V_m , der unter V_r liegt, da g_K noch erhöht ist. V_m nähert sich E_k ($\sim -80mV$), wozu auch eine erhöhte Pumprate der Na^+-K^+ -ATPase⁵ beitragen kann (vgl. [SD07, 46]).

5. Refraktärphase⁶

- *absolute*: Ca. 2 ms. nach Auslösen des Aktionspotenzials sind die Na^+ Kanäle *inaktiviert* und dadurch nicht aktivierbar (vgl. [FE19, 70]). Es ist keine Aktionspotenzialbildung möglich (vgl. [SD07, 46]).

- *relative*: V_m nähert sich weiter V_r an. Nachdem einige Na^+ -Kanäle *deinaktiviert* wurden, ist eine Auslösung des Aktionspotenzials wieder möglich. Allerdings ist die Reizschwelle erhöht (da $V_m < V_r < V_t$) und "die Amplitude des auslösbaren Aktionspotenzials ist reduziert" [FE19, 70].

A.1.1 Übertragungsgeschwindigkeit von Signalen in Nervenzellen

Ist das Axon - die Nervenfasern - **marklos**, wird das Signal kontinuierlich weitergeleitet, und die Leitungsgeschwindigkeit ist eher gering (ca. $1m/s$). Die Leitungsgeschwindigkeit hängt hier direkt von dem Durchmesser der Nervenfasern ab und ist proportional zur Wurzel des Faseradius. Die marklosen Riesenaxone des Tintenfisches erreichen deshalb aufgrund ihrer Größe Leitungsgeschwindigkeiten von bis zu $20m/s$ (vgl. [Jon19, 79]).

¹ den es nicht erreicht. Erreicht werden Werte zwischen 0 und $+40mV$ (vgl. [FE19, 69])

² vgl. [BCP18, 105] und [SD07, 47, Tafel 2.3 (A.2)]

³ Die Kaliumleitfähigkeit wird "verzögerter Gleichrichter" genannt, da sie das ursprüngliche Membranpotenzial - mit Verzögerung - wiederherstellt (vgl. [BCP18, 103]).

⁴ auch: "Nachhyperpolarisation" [SD07, 46]

⁵ Die Na^+-K^+ -ATPase "erhält die Ionenkonzentrationsgradienten aufrecht, die für den Fluss von Na^+ und K^+ durch die Kanäle während des Aktionspotenzials erforderlich sind" (vgl. [BCP18, 105]).

⁶ Die Refraktärphase dient u.a. dazu, die Membran vor einer vorzeitigen Neuerregung zu schützen (vgl. [Jon19, 76]). Hochfrequenten Aktionspotenzialsalven von max. $1000/s$ sind aufgrund dieser Eigenschaft möglich (vgl. [SD07, 46] sowie [BCP18, 89]). *Bear et al.* stellen fest: "Die Frequenz der Aktionspotenziale ist ein Maß für die Stärke des depolarisierenden Stroms" [BCP18, 89] - je stärker der Reiz, desto mehr Aktionspotenziale werden nacheinander abgefeuert (vgl. [BCP18, 90, Abb. 4.3]).

Im Gegensatz zu marklosen Nervenfasern erreichen **myelinisierte Axone** eine höhere Leitungsgeschwindigkeit. Sie sind gegenüber ihrer Umgebung durch **Myelinscheiden**⁷ besser isoliert (vgl. [SD07, 48]). Dadurch wird der Stromfluss verstärkt und die Erregungsleitung erhöht sich. Diese Isolierschicht ist “segmentiert“: Myelinscheiden sind in Abständen von $0,2 - 2\text{mm}$ durch sog. *Ranvier-Schnürringe* unterbrochen. Deren Membran besitzt spannungsabhängige Na^+ -Kanäle, die das ankommende elektrische Signal durch Depolarisierung der einzelnen Segmentabschnitte kontinuierlich weiterleiten - es wird an den Membranabschnitten jeweils ein neues Aktionspotenzial gebildet (vgl. [SD07, 48]). Die Weiterleitung in myelinisierten Axonen bezeichnet man deshalb als **saltatorische** (sprunghafte) Erregungsleitung (vgl. [BCP18, 109 f.]).

A.2 ATPasen

Kurzform für *Adenosintriphosphasen*: Enzyme, die ATP in ADP und Phosphat aufspalten (vgl. [SD07, 26]).

A.3 Dendriten

von “ $\delta\acute{\epsilon}\nu\delta\rho\nu$ (dendrón)“ (altgriechisch): Baum; einzelne selten länger als 2mm (vgl. [BCP18, 28]). Längere Dendriten finden sich an den kortikalen Pyramidenzellen mit einer Länge von 1cm (vgl. [Eil19, 58]).

A.4 Diffusion

unter der Diffusion (“*diffundere*“ (lat.): zerstreuen, ausbreiten) von Molekülen versteht man ihr Bestreben, entlang eines Konzentrationsgradienten (auch: Konzentrationsgefälles) einen Ausgleich der Konzentrationsunterschiede zu erreichen. Moleküle in hoher Konzentration diffundieren dann in die Bereiche mit niedriger Konzentration: In den in Abschnitt 2.1 betrachteten Beispielen diffundieren bspw. K^+ -Ionen, bis ein *Gleichgewichtspotenzial* erreicht ist.

A.5 Effektoren

“*efficere*“ (lat.): bewirken, hervorbringen. Effektoren können wir uns als Endglied der Signalübertragung vorstellen, auch wenn hier wieder interzelluläre Vorgänge stattfinden. Vgl. “neuromuskuläre Endplatte“ in [BCP18, 127, Abs. 3].

⁷ Membranschichten, die sich bis zu 100 mal um das Axon wickeln (vgl. [Jon19, 79])

A.6 Expertensysteme

1990 von *Friedrich und Stumptner* als “kommerziell erfolgreichste Teildisziplin der Artificial intelligence“ bezeichnet [FS90, 14], und ebenda beschrieben als:

Ziel der Expertensysteme ist es, dem Anwender Wissen und Fertigkeiten zur Verfügung zu stellen, über die normalerweise nur speziell ausgebildete oder erfahrene Personen (Experten) verfügen. ([FS90, 14])

Im groben besteht ein Expertensystem aus einer domänenspezifischen *Wissensbasis*, auf der ein *Inferenzmotor* zum Finden von Antworten und Schlussfolgerungen operiert. *Russel und Norvig* erklären, dass Expertensysteme

optimale Entscheidungen empfehlen und dabei die Prioritäten des Benutzers sowie die verfügbaren Evidenzen berücksichtigen ([RN12, 737])

A.7 Goldman-Gleichung

Auch: **Goldman-Hodgkin-Katz-Gleichung** (GHK-Gleichung) nach David Eliot Goldman (1910 – 1998), Alan Lloyd Hodgkin(1914 - 1998) und Bernard Katz (1911 - 2003).

Wie wir in Abschnitt 2.1 gesehen haben, liegt V_r zwischen $-70mV$ und $-90mV$. Wie kann man nun schließen, dass das Ruhepotenzial durch die Membranpermeabilität von K^+ bestimmt wird, wenn $V_r = -70mV$, aber $E_K = -80mV$, und die Membran auch noch für andere Ionen wie bspw. Na^+ selektiv permeabel ist⁸? Wäre die Membran nur für K^+ permeabel, so läge V_r sicher bei E_k (vgl. [SD07, 32]): Die Nernst-Gleichung kann deshalb nur zur Bestimmung des Membranpotenzials genutzt werden, wenn die Membran nur für ein Ion permeabel ist. Ansonsten ist der erhaltene Wert nur näherungsweise zu verstehen (vgl. [FE19, 67]).

Das Membranpotenzial kann auch unter Berücksichtigung mehrerer Ionen bestimmt werden: Ionenkanäle unterstützen einen *passiven Transport* der Ionen zwischen EZF und IZF *entlang* ihres Konzentrationsgefälles, während Ionenpumpen, die *entgegen* des Konzentrationsgefälles arbeiten, *aktiv transportieren* (vgl. [Fro19, 30])⁹. Ionenpumpen sind für die Ionenkonzentrationsgradienten und deren Aufrechterhaltung verantwortlich (vgl. [BCP18, 76])¹⁰. Um V_m zu berechnen müssen die Ionen mitberücksichtigt werden, für die die Membran permeabel ist. Dazu kann die **Goldman-Gleichung** genutzt werden¹¹.

⁸ vgl. [BCP18, 77] sowie [SD07, 44]: “Warum ist E_m weniger negativ als E_K ?”

⁹ hierfür wird metabolische Energie verbraucht (vgl. [Fro19, 31]).

¹⁰ es wird ein nicht unwesentlicher Teil von Energie zur Aufrechterhaltung dieser Gradienten verbraucht. Die Natrium-Kalium-Pumpe verbraucht laut *Bear et al.* etwa 70 % der ATP-Menge, die das Gehirn benötigt (vgl. [BCP18, 76]).

¹¹ *Silbernagl und Despopoulos* nutzen für die Bestimmung von V_m die fraktionelle Leitfähigkeit der involvierten Ionen und rechnen $V_r = E_K \times f_K + E_{Na} \times f_{Na} + E_{Cl} \times f_{Cl}$ (vgl. [SD07, 32, 1.21])

Kandel et al. stellen hierzu fest, dass eine hohe Konzentration eines einzelnen Ions zusammen mit einer hohen Membranpermeabilität für dieses Ion auch einen größeren Beitrag für V_r leistet:

when permeability to one ion is exceptionally high, the Goldman equation reduces to the Nernst equation for that ion. ([Kan+13, 135])

Fakler und Eilers weisen darauf hin,

dass die Permeabilitäten in komplizierter Weise von der Membranspannung und den Ionenkonzentrationen [...] abhängen und sich meist nur näherungsweise bestimmen lassen. ([FE19, 67])

A.8 KI Winter

Der Begriff “KI Winter“ wird in der Literatur mit unterschiedlichen Perioden während der Forschung und Förderung von KI in Zusammenhang gebracht. In dem Kontext des im Abschnitt 4.1 erwähnten Lighthill Reports bezieht sich der Begriff auf die Periode nach 1973:

Lighthill’s report provoked a massive loss of confidence in AI by the academic establishment in the UK (and to a lesser extent in the US). It persisted for a decade - the so-called ‘AI Winter’. (Artificial intelligence at edinburgh university: A perspective¹²)

Russell und Norvig beziehen sich auf einen Zeitraum um/ nach 1988, in dem nach einer Phase von Investitionen in Milliardenhöhe in den Forschungszweig “viele Unternehmen verschwanden, weil sie ihre außergewöhnlichen Versprechungen nicht halten konnten.“ [RN12, 48]. Auf gleichen Zeitraum bezieht sich *McCorduck* in [McC04, 432]; vgl. hierzu auch:

Dozens of expert systems companies and AI-focused hardware manufacturers failed *en masse* as hype turned to disillusionment. ([Gar19, 656; Hervorhebung i.O.]

A.9 LeNet

LeCun et al. weisen darauf hin, dass LeNet eine Erweiterung der von LeCun beschriebenen Architektur in [LeC89] ist (vgl. [LeC+89, 544]).

Spätere Iterationen von *LeNet* durch *LeNet-4* und *LeNet-5*. Die Architektur von *LeNet-5* wird später angemessen mit der in Mitte der 90er Jahre zur Verfügung stehenden Technologie skaliert:

In 1989 a recognizer as complex as LeNet-5 would have required several weeks’ of training, and more data than was available, and was therefore not even considered. ([LeC+98, 15])

¹² <https://www.inf.ed.ac.uk/about/AIhistory.html>, abgerufen 31.08.2023

Eingabedaten für Lenet-5 sind ein 32×32 Pixel großes Bild, außerdem besitzt das Netz 7 Schichten (ohne Eingabeschicht) (*LeNet-1*: 3 verborgene Schichten (vgl. [LeC+89, 544])). Eine vollständige Beschreibung der eindrucksvollen Architektur findet sich in [LeC+98, 7 f.]

A.10 Lernrate

Die **Lernrate** (η , in der Literatur auch α) ist der Koeffizient für die Gewichtsadjustierungen in einem künstlichen Neuron.

Wie in Abschnitt 3.2.4 gesehen, berechnet die Lernregel die Gewichte auf Basis des *Fehlers*, also der Differenz von erwarteter Ausgabe und tatsächlicher Ausgabe: Ist in den Beispielen der Fehler -1 , werden die Gewichte verringert, ist der Fehler 1 , werden die Gewichte erhöht. Die Lernrate η ist der Koeffizient für die Gewichtsadjustierung, und wird auch *Schrittweite* genannt (vgl. [GBC18, 93] und [RN12, 172]).

Üblicherweise liegt η zwischen 0 und 1¹³.

A.11 Neurotransmitter und ihre Rezeptoren

Die Diffusion der Vesikel in den synaptischen Spalt dauert ca. 10 – 100 μs , danach binden sich die Transmitter an die Rezeptoren der postsynaptischen Zelle (vgl. [HS19a, 96]), die das “interzelluläre chemische Signal [...] in ein intrazelluläres Signal (eine Änderung des Membranpotenzials oder eine chemische Veränderung) in der postsynaptischen Zelle umwandeln“ [BCP18, 123].

Rezeptoren können hierbei *ionotrop* oder *metabotrop* sein: Inotrope Rezeptoren sind gleichzeitig auch Ionenkanäle und aktivieren sich, wenn ein bestimmter Transmitter an sie bindet (vgl. [HS19b, 109]).

Metabotrope Rezeptoren lösen intrazelluläre Stoffwechselfvorgänge aus, und weitere Botenstoffe (sog. **second messenger**) können dann für eine Aktivierung von Ionenkanälen verantwortlich sein (vgl. [RK18, 134]).

Als Neurotransmitter spielen im zentralen Nervensystem vor allem die Aminosäuren *Glutamat* (Glu, erregend), *Gamma-Aminobuttersäure* (GABA, hemmend)¹⁴ und Glycin (Gly, hemmend) eine Rolle. An neuromuskulären Endplatten vermittelt Amin *Acetylcholin* (ACh, erregend)¹⁵.

Schnelle Formen der synaptischen Übertragung dauern zwischen 2 und 100 ms , langsame Übertragungen einige 100 Millisekunden bis Minuten (vgl. [BCP18, 129

¹³ vgl. [Fau94, 61]; außerdem: “ α [η] ist eine kleine Konstante“ [RN12, 172]. Gleiche Stelle beleuchtet das Für und Wider kleiner und großer Werte von η ; *Salomon* weist in [Sal90, 173] darauf hin, dass eine geeignete Lernrate auch von der Aufgabenstellung abhängt.

¹⁴ GABA: Gamma-Aminobuttersäure. Um den GABA-Spiegel zu erhöhen, kann bspw. *Gabapentin* verabreicht werden (<https://www.aerzteblatt.de/archiv/20049/Neuropathien-Gabapentin-brems-t-ueberaktive-Neurone>, abgerufen 01.08.2023): Es ist bei der Behandlung von Anfallsleiden wie der Epilepsie sowie bei Nervenschmerzen (Neuropathien) wirksam (https://www.gelbe-liste.de/wirkstoffe/Gabapentin_21579, abgerufen 01.08.2023)

¹⁵ zur Entdeckung von ACh siehe Anhang B.5

f.]). Transmitter werden oft zusammen mit Co-Transmittern ausgeschüttet, die die Erregungsübertragung modulieren (vgl. [SD07, 52]).

Ob ein exzitatorischer oder inhibitorischer Transmitter auch dieselbe Wirkung in der postsynaptische Zelle hervorruft, entscheidet sich bei den Rezeptoren (vgl. [HS19b, 109]). Als Beispiel sei das bereits oben erwähnte ACh genannt, das die Kontraktion des Herzens verlangsamt, bei der Skelettmuskulatur jedoch eine schnelle Depolarisation der Muskelfasern bewirkt (vgl. [BCP18, 137]).

Erregende und hemmende Synapsen kann man auch an ihrer Struktur erkennen: Asymmetrische oder *Gray-Typ-I-Synapsen* sind auf der postsynaptischen Seite dicker als auf der präsynaptischen, bei gleicher Dimension der Membrandifferenzierungen spricht man von *Gray-Typ-II-Synapsen*. Typ-I sind in der Regel exzitatorisch, Typ-II inhibitorisch (vgl. [BCP18, 127 u. 147]).

A.12 Perzeptron

The Perceptron - A perceiving and recognizing automaton

Das Forschungsprojekt “*Perceiving and Recognizing Automaton*“ beschreibt einen Apparat, der mittels einer Kamera geometrische Figuren erkennen und zuordnen kann (vgl. [Ros57, 3]). Die Funktionen simuliert Rosenblatt zunächst auf einem IBM 704 Rechner¹⁶, bevor die Hardware Anfang der 60er Jahre als *Mark 1 Perceptron* gebaut wird: 400 Cadmiumsulfid-Photozellen auf einem 20x20 großen Raster angeordnet - dem **S-System** (**S** = *Sensory*) - leiten Signale an das **A-System** (**A** = *Association*); dort werden sie registriert und ausgewertet, und schlussendlich über das **R-System** (**R** = *Response*) ausgegeben (vgl. [Ros57, 4 f.], [Ros58, 389 f.] sowie [Bis06, 193, “Frank Rosenblatt] und [Bis06, 196, Figure 4.8]). Dabei lernt die Maschine im ersten Schritt durch die Unterstützung der Ingenieure, wie gegebene Formen zu interpretieren sind: Für aktivierte Photozellen wird die erwartete Ausgabe manuell festgelegt. Die Verbindungen zwischen den **S**-, **A**- und **R**-Units erinnert nicht nur von der Namensgebung her an biologische Neuronen, auch deren Struktur und Verschaltung wird hier als Vorbild genommen (vgl. [Ros62, 4]).

Die **S**-Units konnten sowohl hemmende als auch erregende Signale in das **A**-System einspeisen. Darüber hinaus war das **R**-System in der Lage, über Rückkopplungen hemmende Signale an das **A**-System zu senden: Damit sollte verhindert werden, dass weitere **R**-Units aktiviert werden, die sich mit den bereits aktivierten Units gegenseitig ausschließen (vgl. [Ros57, 4 f.]).

¹⁶ ausführlich beschrieben in [Ros60]

A.13 Shunting Inhibition

Kurzschlusshemmung bei Neuronen

Wenn ein Transmitter einer inhibitorischen Synapse Cl^- -Kanäle aktiviert, verschiebt sich das Membranpotenzial in Richtung $E_{Cl} \sim -65mV$. Ist das Membranpotenzial der Zelle $> -65mV$, kann ein hyperpolarisierendes IPSP ausgelöst werden. Wenn allerdings bereits $V_m = E_{Cl}$ ist¹⁷, wird kein "sichtbares" IPSP ausgelöst (vgl. [BCP18, 145 f.]) - die Cl^- -Kanäle sind offen, aber es findet keine Nettoionenbewegung statt. Eine Hemmung der Zelle kann aber trotzdem noch stattfinden, wenn bspw. ein distal liegende exzitatorische Synapse positive Ladungsströme verursacht, und die inhibitorische Synapse proximal zum Soma liegt, in Stromrichtung ausgehend von der exzitatorischen Synapse.

Der positive Ladungsstrom kommt dann an der inhibitorischen Synapse mit den geöffneten Cl^- -Kanälen an, und die Depolarisation durch das EPSP wird durch den Einfluss von Cl^- -Ionen (die V_m wieder auf $V_{rev} = E_{Cl}$ bringen wollen) mitigiert.

Aus diesem Grund sind inhibitorische Synapsen oft proximal zum Soma zu finden (vgl. [Kan+13, 231]).

A.14 Soma

von "σῶμα (sōma)" (altgriechisch): Körper; auch *Perikaryon* [RK18, 58]. Bezeichnet den Zellkörper und das Stoffwechsellzentrum des Neurons mit der eine Größe von ca. 20 μm [BCP18, 29]. Zum Vergleich: ein menschliches Haar hat einen Durchmesser von ca. 70 μm , kleine Bakterien bis zu 20 μm .

A.15 Umkehrpotenzial

Wie wir bei der Entstehung des Aktionspotenzials (siehe Abschnitt 2.2) gesehen haben, transportieren spannungsgesteuerte Ionenkanäle stets entlang des elektrochemischen Gradienten (vgl. [Fak19, 39]). Im Folgenden sei die Membran zur Vereinfachung nur permeabel für ein Ion, es gelte weiterhin $V_m < E_{Ion}$: Die Stromrichtung für das Ion ist einwärts IZF. Ist die Triebkraft positiv wegen $V_m > E_{Ion}$, strömt das Ion auswärts EZF. Wenn allerdings

$$V_m - E_{Ion} = 0 \tag{A.1}$$

dann gilt:

$$V_m = E_{Ion} \tag{A.2}$$

¹⁷ "häufig bei für Chloridleitfähigen GABA Rezeptoren" [HS19a, 100]

In diesem Fall entspricht die Membranspannung dem Gleichgewichtspotenzial des Ions - es findet keine Nettoionenbewegung statt.

Der Wert für V_m , bei dem die Differenz zwischen V_m und E_{Ion} bei 0 liegt, wird das *Umkehrpotenzial* V_{rev} genannt, weil hier ein Vorzeichenwechsel stattfindet: Je nachdem, in welche Richtung sich V_m ändert, ändert sich auch die Richtung des Stroms. Für ionenspezifische Kanäle entspricht das *Umkehrpotenzial* ihrem *Gleichgewichtspotenzial* (vgl. [McC13, 95 f.]).

Sobald eine Membran permeabel für mehr als ein spezifisches Ion ist, muss nach *Kandel et al.* die relative Leitfähigkeit der Membran für diese Ionen sowie deren Gleichgewichtspotenziale zur Bestimmung des Umkehrpotenzials berücksichtigt werden (vgl. [Kan+13, 196, Box 9–1]). Für die ACh-Rezeptoren an der motorischen Endplatte¹⁸, die für Na^+ und K^+ permeabel sind, folgt damit, dass die Summe ihrer Ströme am Umkehrpotenzial $V_{rev} = 0$ sein muss:

$$I_{Ka} + I_{Na} = 0 \quad (\text{A.3})$$

Da der Membranstrom I_{Ion} gleich dem Produkt der Membranleitfähigkeit und der elektrochemischen Triebkraft für dieses Ion ist (vgl. [BCP18, 93]), also

$$I_{Ion} = g_{Ion}(V_m - E_{Ion}) \quad (\text{A.4})$$

liefert Ersetzen die Gleichung

$$g_{Na}(V_m - E_{Na}) = g_K(V_m - E_K) \quad (\text{A.5})$$

Da am Umkehrpotenzial $V_{rev} = V_m$ gilt, können wir V_m durch E_{rev} ersetzen und danach auflösen, was zur folgenden Gleichung führt (vgl. [Kan+13, 196, Box 9-1]):

$$E_{rev} = \frac{E_{Na} * (g_{Na}/g_K) + E_K}{(g_{Na}/g_K) + 1} \quad (\text{A.6})$$

Die ACh-Rezeptoren haben an der Endplatte eine relative Leitfähigkeit für $Na^+ : K^+$ von 1,8, für die Gleichgewichtspotenziale gilt $E_{Na} = +55mV$ und $E_K = -100mV$; somit folgt nach Einsetzen $E_{rev} = 0mV$ (vgl. [HS19a, 100]). Wenn nun das Membranpotenzial vor der Einwirkung von ACh $< 0mV$ ist, bewirkt ein Öffnen der Ionenkanäle einen Strom einwärts IZF, um V_m auf 0 zu bringen (Depolarisation); im umgekehrten Fall $V_m > 0$ fließen Ionen auswärts zu EZF, und es findet eine Hyperpolarisation statt (vgl. [BCP18, 136, Exkurs 5.4]).

¹⁸ chemische Synapse, die Erregungen an eine Muskelfaser weiterleitet (vgl. [SD07, 56])

B

Personen

B.1 Bernstein, Julius

Die Entdeckung bioelektrischer Ströme

Im Jahr 1912 veröffentlichte der deutsche Physiologe *Julius Bernstein* (1839-1917) Erkenntnisse seiner elektrophysiologischen Studien in dem Buch “*Elektrobiologie - Die Lehre von den elektrischen Vorgängen im Organismus auf moderner Grundlage dargestellt*“. In dem Kapitel “Membrantheorie“ in [Ber12] greift er das Postulat seiner Arbeit “Untersuchungen zur Thermodynamik der bioelektrischen Ströme“ aus [Ber02] auf.

Bereits 1902 von ihm veröffentlicht, lieferte die Theorie die erste plausible physikalisch-chemische Erklärung bioelektrischer Phänomene (vgl. [Sey06, 5]): In Zellen befindet sich eine elektrolytische Flüssigkeit, die von einer für bestimmte Ionen permeablen Membran umgeben ist, die eine Potenzialdifferenz aufweist (vgl. [Ber12, 92 f.]):

Denken wir uns, dass diese Elektrolyte aus dem Querschnitt der Fibrillen ungehindert in die umgebende Flüssigkeit diffundieren, während sie am Längsschnitt durch die lebende Sarkoplasmahaut daran gehindert werden, weil sie für ein Ion derselben, z. B. für das Anion (PO_4^- u. s. w.), mehr oder weniger impermeabel ist, so entstünde auf der Oberfläche der Fibrille eine elektrische Doppelschicht, welche nach innen negative, nach aussen positive Spannung besitzen würde. ([Ber02, 542])

Knapp 34 Jahre vor seiner Membrantheorie vermutet Bernstein bereits die Existenz eines Aktionspotenzials (vgl. [Sch83, 168]). Seine Forschungsarbeit über elektrische Muskel-Nerven-Aktivitäten in [Ber68] schließt er mit den Worten:

Ich werde im Folgenden diese Welle mit dem Namen ‘Reizwelle’ bezeichnen, weil derjenige Reiz, durch den die Nervenfasern im Centrum Empfindung, im Muskel Zuckung erzeugt, auf das Innigste mit dieser Welle verknüpft ist. [...] Da wir nachgewiesen haben, dass die Reizwelle mit derselben Geschwindigkeit sich fortpflanzt als die Erregung, so können wir die wohlberechtigte Annahme machen, dass die Reizwelle Nichts anderes ist als das Bild des im Nerven ablaufenden Erregungsvorganges. ([Ber68, 198 f.])

In den darauffolgenden Jahren und Jahrzehnten haben zahlreiche Forschungen die Theorie bestätigt: Wir wissen heute, dass die ungleiche Ionenverteilung in der intrazellulären und extrazellulären Flüssigkeit zusammen mit der Permeabilität der Membran für die Entstehung eines Membranpotenzials verantwortlich ist¹].

B.2 Hebb, Donald

Cells that fire together, wire together: Synaptische Plastizität

Donald Hebb (1904 - 1985), gebürtiger Kanadier, Sohn eines Ärzte-Elternpaares und 1965 nominiert für den Nobelpreis, geht als junger Mann einer Karriere als Schriftsteller nach (vgl. [BM03, 1013 f.]).

Er studiert Englisch als Hauptfach und macht 1924 seinen Bachelor², doch die Schriften Freuds, mit denen er sich nach seinem Abschluss beschäftigt, wecken in ihm den Wunsch, sich tiefer mit Psychologie zu beschäftigen (vgl. [BM03, 1013]): An der McGill Universität in Montreal³ macht er 1932 seinen Master darin⁴, und leitet dort 16 Jahre später als Professor die Fakultät für Psychologie (vgl. [Coo05, 853]).

Seine Faszination darüber, wie das Gehirn lernt, Informationen verarbeitet und speichert⁵ wird Bestandteil seiner Forschungsarbeit: 1949 veröffentlicht er das Buch **“The organization of Behavior: A Neuropsychological Theory“** [Heb49]; seine darin formulierten Postulate⁶ liefern einen wichtigen Beitrag für die Neurowissenschaften. Als hätte sein Buch eine Art Golgräberstimmung in der Psychologie ausgelöst, schreibt *Klein* hierzu:

“It attracted many brilliant scientists into psychology, made McGill University a North American mecca for scientists interested in brain mechanisms of behaviour, led to many important discoveries, and steered contemporary psychology onto a more fruitful path.“ ([Kle99, 1])

Oft zitiert wird seine Idee bzgl. synaptischer Verstärkung, was heute als **Hebb-sche Synapse** bekannt ist⁷:

When an axion of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change

¹ siehe hierzu Kapitel 2

² an der Dalhousie Universität: <https://dal.ca> (abgerufen 17.08.2023), vgl. [Coo05, 852]

³ McGillUniversität, Montreal, Quebec (Kanada): <https://www.mcgill.ca/neuro/about/donald-hebb-phd> (abgerufen 16.08.2023)

⁴ zu der Zeit studierte er in Teilzeit an der McGill Universität: “as a part time graduate student“ [Kle99, 1]. Seine Master-Arbeit schrieb er aufgrund einer Erkrankung im Bett (vgl. [BM03, 1014])

⁵ vgl. Eintrag über Hebb in [Str01, 298 f.]

⁶ es sind “three pivotal postulates“ [Kle99, 2]

⁷ siehe [AR88, 43] sowie als Zitat: *a*: “any two cells or systems of cells that are repeatedly active at the same time will tend to become ‘associated’, so that activity in one facilitates activity in the other.“ [Heb88, 52] sowie *b*: “A series of such events [Aktivierung von cell-assemblies] constitutes a ‘phase sequence’ - the thought process“. [Heb88, 48]; die Aussage findet sich im Original bereits in [Heb49, xi-xix, “Introduction“]

takes place in one or both cells such that A's efficiency, as one of the cells firing B, is increased. ([Heb88, 50; Hervorhebung i.O.]

Derartige Veränderungen synaptischer Verbindungen wird als **Hebbsche Lernregel** (vgl. [BCP18, 985]) bezeichnet. Das damit verbundene geflügelte Wort *Cells that fire together, wire together* beschreibt die Hypothese bildhaft⁸. Seine Idee der **“Cell Assembly“** schließt daran an: Damit sind Verbände von Neuronen gemeint, die miteinander verschaltet sind, und deren Verbindungen durch das Hebbsche Lernen so sehr verstärkt sind, das die Aktivierung einzelner Zellen in diesen Verbänden ausreicht, das alle Zellen aktiviert werden (vgl. [BCP18, 907 f.]).

Hebbs Theorien gelten durch die Forschung als bestätigt⁹, und mit der von Hebb formulierten synaptische Plastizität wurde auch eine Idee für lernende künstliche neuronale Netze geliefert:

[Hebb] laid the foundation for neoconnectionism which seeks to explain cognitive processes in terms of connections between assemblies of real or artificial neurons. ([Kle99, 2])

B.3 Hodgkin, Alan und Huxley, Andrew

Für die Erforschungen der Ionen-Mechanismen, die bei der Erregung und Hemmung von Nervenzellmembranen beteiligt sind, erhalten *Alan L. Hodgkin* (1914 - 1998) und *Andrew F. Huxley* (1917 - 2012) im Jahr 1963 den Nobelpreis für Physiologie oder Medizin, zusammen mit *John Carew Eccles* (1903 – 1997) (vgl. [Gle09]).

Knapp 80 Jahre nach Bernsteins Beschreibung der “negativen Schwankungen“ und 50 Jahre nach seiner Membrantheorie¹⁰ führen Hodgkin und Huxley Experimente an Riesenaxonen von Tintenfischen durch, die zu einem mathematischen Modell des Membranverhaltens führen.

Einige Jahre vorher - 1939 - können *Curtis* und *Cole* bei ähnlichen Versuchsanordnungen während der Untersuchung von Membraneigenschaften in Folge eines Aktionspotenzials einen sprunghaften Anstieg der Leitfähigkeit der Membran feststellen (vgl. [CC39, 669]).

Die Experimente wurden an Riesenaxomen von Tintenfischen durchgeführt, die sich für die damals zur Verfügung stehende Technologie mit einer Dicke von ca. *1mm* besser eigneten als die im direkten Vergleich geradezu unfassbar winzigen Axone von menschlichen Nervenzellen mit einem Durchmesser von $\sim 1 \mu\text{m}$ (vgl. [Jon19, 79]):

⁸ Zumindest in [Heb49] soll sich kein solches Zitat finden. *Keyzers und Gazzola* stellen hierzu fest: “This mnemonic phrase was first introduced by Carla Shatz in an article for the Scientific American aimed at lay public“ [KG14, 2, Fussnotenmarker entfernt] und meint damit in [Sha92] den Satz “In a sense, then, cells that fire together wire together.“ [Sha92, 64]. Die Autoren des Eintrags https://en.wikipedia.org/wiki/Hebbian_theory#cite_ref-2 (abgerufen 16.08.2023) hingegen schreiben den Ursprung [LS92] von *Löwel und Singer* zu: “neurons wire together if they fire together.“ [LS92, 211]

⁹ vgl. [Flo19, 833]. *Bear et al.* verweisen in [BCP18, 875, Exkurs 23.5] auf [CL78]

¹⁰ siehe hierzu Anhang B.1

The large diameter of the axon, 0.5 mm or more, makes it particularly favorable material [...] ([CC39, 650])

Hodgkin und Huxley weisen später nach, dass der Aufstrich des Aktionspotenzials mit einer Zunahme von g_{Na} und dem Einstrom von Na^+ zusammenhängt, und weiter, dass die Repolarisation der Zellmembran mit einer Zunahme von g_K und einem Ausstrom von K^+ zusammenhängt¹¹. Sie stellen Hypothesen unter anderem zu dem Schwellenwert

The curves in Figs. 12 and 21 show that the theoretical 'membrane' has a definite threshold when stimulated by a sudden displacement of membrane potential. ([HH52, 535])

und der Refraktärzeit auf:

According to our theory, there are two changes resulting from the depolarization during a spike which make the membrane unable to respond to another stimulus until a certain time has elapsed. These are 'inactivation', which reduces the level to which the sodium conductance can be raised by a depolarization, and the delayed rise in potassium conductance, which tends to hold the membrane potential near to the equilibrium value for potassium ions. ([HH52, 532])

Außerdem formulieren sie ein mathematisches Modell auf Basis des beobachteten Membranverhaltens (vgl. [HH52, 501, Fig. 1.]), heute bekannt als das **Hodgkin-Huxley-Modell** (vgl. [Koc98, 144]). Die Relevanz des Modells fassen *Kandel et al.* zusammen:

More than a half-century later the Hodgkin-Huxley model stands as the most successful quantitative model in neural science if not in all of biology. ([Kan+13, 156])

B.4 Hubel, David und Wiesel, Torsten

David H. Hubel (1926 - 2013) und *Torsten N. Wiesel* erhalten 1981 den Nobelpreis für Physiologie oder Medizin für ihre Entdeckungen bzgl. der Informationsverarbeitung im Seewahrnehmungssystem¹². In [HW62] (1962) zeigen sie, dass im Wahrnehmungssystem "einfache" und "komplexe" Neuronen existieren, die visuelle Informationen unterschiedlich verarbeiten (vgl. [Wur09, 2819])

B.5 Loewi, Otto

Die Entdeckung des Vagusstoff

Otto Loewi (1873 - 1961) erhält 1936 den Nobelpreis für Medizin für seine Forschungen an der chemischen Übertragung von Nervenimpulsen.

¹¹ vgl. [Jon19, 75], außerdem [BCP18, 96] sowie [HH52, 530 Fig. 17]

¹² <https://www.nobelprize.org/prizes/medicine/1981/summary>, abgerufen 05.09.2023

1921 kann er bei einem Experiment mit Froschpräparaten zeigen, dass die Infusionslösung, die für vorher am Vagusnerv bewusst stimulierte Froschherzen genutzt wurde, den Vagusnerv von nachträglich mit dieser Lösung behandelte Froschherzen stimulieren kann.

Loewi vermutet in der Lösung eine Substanz, die er "Vagusstoff" nennt. Henry Dale (1875 - 1968) - mit dem sich Loewi den Nobelpreis teilen wird - kann diesen Stoff später als *Acetylcholin* identifizieren, ein exzitatorischer Neurotransmitter¹³, der bei chemischen Synapsen als Botenstoff zur Signalübertragung beteiligt ist¹⁴.

B.6 McCulloch, Warren und Pitts, Walter

Von der Philosophie zur Neurowissenschaft

Warren McCulloch (1898 - 1969), studierter Psychologe, Philosoph und Neurophysiologe, wird früh in seiner Karriere¹⁵ von grundlegenden philosophische Fragen umtrieben (vgl. [Arb00, 1]):

What is a number, that a man may know it, and a man that he may know a number? ([McC16, 2])

Die Suche nach einer Antwort auf diese und ähnliche Fragen¹⁶ soll ihn Zeit seines Lebens beschäftigen und Bestandteil seiner Arbeit werden (vgl. [Abr02, 4 f.]): Es drängt ihn in die Neurowissenschaften¹⁷.

Um psychische Ereignisse auf das wesentlichste zu reduzieren und besser erklärbar zu machen, formuliert er das "*psychon*", ein einzelnes psychisches Ereignis, das u.a. folgende Eigenschaften hat:

1. Es tritt auf, oder es tritt nicht auf.
2. Wenn es auftritt, dann nur, weil es von einem vorhergehenden Ereignis ausgelöst wurde.
3. Es soll ein nachgehendes "psychon" auslösen können¹⁸.

Die Ähnlichkeiten zu dem **Alles-oder-Nichts-Prinzip** von Nervenzellen zur Formalisierung seines Kalküls berücksichtigt er ab 1929:

¹³ "Acetylcholin wirkt erregend an der motorischen Endplatte, aber hemmend an den Schrittmacherzellen des Herzens" [HS19b, 105].

¹⁴ vgl. [BCP18, 119, Exkurs 5.1]

¹⁵ die im folgenden zitierte Frage wird von der Quelle auf das Jahr 1917 datiert, McCullochs erstem College-Jahr in Haverford, Pennsylvania

¹⁶ "In addition to the work in functional neuroanatomy, McCulloch was continuing to pursue philosophical questions, such as 'What could be the logic of the brain?'" [Arb00, 3]

¹⁷ siehe hierzu "The Logical Structure of Mind: An Inquiry into the Philosophical Foundation of Psychology and Psychiatry", S. iii; <https://ntrs.nasa.gov/citations/19650017787>, abgerufen 19.09.2023

¹⁸ Charakteristika der zweiwertigen Aussagenlogik sind hier bereits erkennbar (vgl. [Abr02, 7]). Eine Form dieser zweiwertigen Logik wird später von dem französischen Psychoanalytiker Jacques Lacan (1901 - 1981) in seine Theorien einbezogen (vgl. [Liu10, 317]).

it dawned on me that these events might be regarded as the all-or-none-impulses of neurons [...]. I began to try to formulate a proper calculus for these events by subscripting symbols for propositions in some sort of calculus of propositions ([McC16, 6])

Knapp 14 Jahre später wird McCulloch in seinem Papier das “psychon“ mit der Aktivität eines einzelnen Neurons gleichsetzen:

[...] a psychon can be no less than the activity of a single neuron. ([MP43, 114])

Doch vorerst stagniert seine Arbeit, und erst, als er nach seiner Berufung an das Neuropsychiatrische Institut der Universität von Illinois (Chicago) im Jahr 1941 *Walter Pitts* (1923 - 1969) kennenlernt, ist er in der Lage, den Kalkül zu vervollständigen.

Pitts besucht dort Kurse bei dem Philosophen und Logiker *Rudolf Carnap* (1891 - 1970), außerdem bei dem Biomathematiker¹⁹ *Nicolas Rashevsky* (1899 - 1972) (vgl. [Pic04, 184]), der in seiner Arbeit mathematische Beschreibungen für die Modellierung von Funktionen von Nerven und Nervennetzen verwendet (vgl. [Abr02, 13])²⁰ sowie ab 1939 die Zeitschrift “Bulletin of Mathematical Biophysics“ veröffentlicht (vgl. [Abr02, 16]): In dieser wird später [MP43] erstmalig publiziert.

Der knapp 19-jährige Pitts gilt als brillianter Mathematiker²¹ und später als “the intelligence behind [...] McCulloch“: Gemeinsam veröffentlichen sie nur wenige Monaten nach ihrem Kennenlernen das “logical calculus“-Papier (vgl. [AR98, 104]).

B.7 Nernst, Walther

Walther Nernst (1864 - 1941) war ein deutscher Physiker und Chemiker und gehört zu den Begründern der physikalischen Chemie. Er formulierte das **Nernst-Theorem**, auch bekannt als **Dritter Hauptsatz der Thermodynamik**. 1920 erhält er den Chemie-Nobelpreis in Anerkennung seiner Arbeit auf dem Gebiet der Thermochemie.

¹⁹ “Biomathematik“ ist ein Teilgebiet der angewandten Mathematik und die mathematische Ausrichtung der theoretischen Biologie, die sich mit Modell- und Theoriebildung zur Beschreibung biologischer Zusammenhänge beschäftigt. (https://de.wikipedia.org/wiki/Theoretische_Biologie, abgerufen 09.08.2023). Teilgebiete der Biomathematik sind u.a. Demographie, Mathematische Ökologie, Epidemiologie, Populationsgenetik

²⁰ “In these works, formal logic and mathematics played a strong role, and thus biology as a discipline was reaching a more mature stage, as it began to incorporate the scientific method of physics, that is, using theoretical analysis and mathematical formulations.“ [Abr02, 7]

²¹ Siehe [Arb00, 4], außerdem [Abr02, 22]: “Pitts, a mathematical prodigy“; *Lettvin* erwähnt in einem Interview mit *Anderson und Rosenfeld*, dass Pitts die *Principia Mathematica* mit 12 Jahren liest (vgl. [AR98, 2])

B.8 Wiener, Norbert

Norber Wiener (1894 - 1964) gilt als Begründer der Kybernetik, die *Küppers* beschreibt als

wirkungsvolle Kommunikation bzw. verlustarmer Daten- und Informationsaustausch, der in der Natur die Überlebensfähigkeit und in der Technik die maschinelle, prozessuale Funktionalität stärkt und dadurch Fehler vermeiden hilft. ([Küp19, 2])

Wiener und McCulloch (s. B.6) lernen sich im Winter 1942 kennen (vgl. [Mas90, 218]). Unter dem Vorsitz von McCulloch wird auf den sog. “Macy-Konferenzen“²² zwischen 1946 und 1953 Biologie und Technologie diskutiert (vgl. [Arb00, 5]). Infolgedessen veröffentlicht Wiener 1948 seine Arbeit “Cybernetics: Or Control and Communication in the Animal and the Machine“ [Wie48].

²² <https://www.asc-cybernetics.org/foundations/history/MacySummary.htm>, abgerufen 09.08.2023

Die Paradoxe Hitzeempfindung als MCP-Netz

Modellierung von MCP-Netzen unter Berücksichtigung von Zeiteinheiten

In Abschnitt 3.1 haben wir MCP-Netze betrachtet, bei denen Zeiteinheiten nicht berücksichtigt wurden. Basierend auf dem von McCulloch und Pitts in [MP43] postulierten Neuronen-Modell entwerfen wir im folgenden ein Netz, das die **paradoxe Hitzeempfindung** simuliert: Das Phänomen, das bei kurzzeitiger Kühlung der Haut Hitze empfunden wird (vgl. [Häm+82]):

If a cold object is held to the skin for a moment and removed, a sensation of heat will be felt; if it is applied for a longer time, the sensation will be only of cold, with no preliminary warmth, however transient. ([MP43, 106])

Das nach [MP43, 105, Figure 1 (e)] aufgestellte Modell und die dazugehörige Formel übertragen wir in die von uns verwendete Notation (in Anlehnung an [Fau94, 31 f.]). In der Abbildung C.1 sind zwei Temperaturrezeptoren H_i (Hitze) und K_i (Kälte) dargestellt, außerdem hat das Netz zwei Ausgänge, H_o und K_o . H_o wird zum Zeitpunkt t aktiviert, wenn Hitze empfunden wird, K_o entsprechend bei Kälte.

Die Bedingungen für eine Aktivierung von H_o lauten: Hitze wird signalisiert, wenn entweder der Wärmerezeptor durch Wärme stimuliert **oder** kurzzeitig ein Kälteimpuls von den Kälterezeptoren registriert wird.

Formal dargestellt¹:

$$H_o(t) \equiv H_i(t-1) \vee (K_i(t-3) \wedge \neg K_i(t-2)) \quad (\text{C.1})$$

Die Bedingungen für eine Aktivierung der Zelle K_o , um Kälte zu signalisieren, lautet

$$K_o(t) \equiv K_i(t-2) \wedge K_i(t-1) \quad (\text{C.2})$$

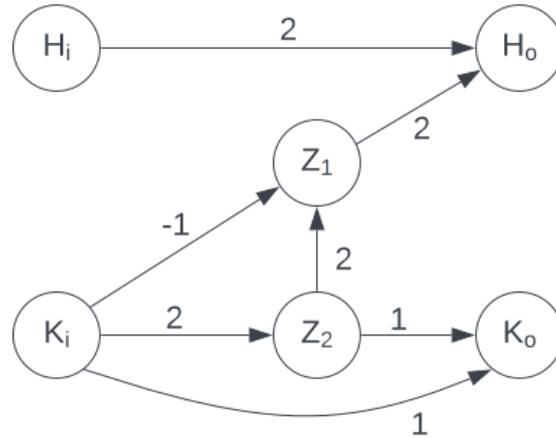


Abb. C.1. MCP-Netz zur Modellierung der paradoxen Hitzeempfindung. (Quelle: in Anlehnung an [Fau94, 31, Figure 1.18])

Zeitabhängige Werte der MCP-Zellen wollen wir nun durch Formeln auszudrücken. Die zu dem Zeitpunkt t bekannten Werte sind bereits in Gleichung C.1 sowie Gleichung C.2 erfasst. Es folgt für $H_o(t)$

$$H_o(t) \equiv Z_1(t-1) \vee H_i(t-1) \quad (\text{C.3})$$

Gleichung C.3 ist abhängig von Z_1 und H_i zum Zeitpunkt $t-1$. $Z_1(t-1)$ ist trivialerweise der Wert, mit dem der Rezeptor stimuliert wurde. Für $Z_1(t-1)$ folgt:

$$Z_1(t-1) \equiv Z_2(t-2) \wedge \neg K_i(t-2) \quad (\text{C.4})$$

Analog zu $H_i(t-1)$ ist $K_i(t-2)$ einfach der Stimulus für den Rezeptor. Deshalb benötigen wir nur noch $Z_2(t-2)$, was das efferente Signal von $K_i(t-3)$ ist:

$$Z_2(t-2) \equiv K_i(t-3) \quad (\text{C.5})$$

Die Abfolge der Aktivierung von K_o bzw. H_o kann tabellarisch auf Grundlage der Gleichungen dargestellt werden. Hierbei ist der Inhalt der Tabellenzellen (bis auf die Spalte t) $f(g(X))$ (siehe Gleichung 3.3), sofern der Wert zu dem entsprechenden Zeitpunkt bekannt ist (vgl. [Fau94, 32 f.]). Wir beginnen mit der Stimulierung von K_i durch zwei aufeinanderfolgende Kälteimpulse (Abbildung C.2). Tabelle C.1 stellt die jeweiligen Schritte dar: Es wird erwartet, dass K_o feuert. Tabelle C.2 beschreibt im Gegenzug die erwartete Ausgabe ‘‘Wärme‘‘ für einen Stimulus von H_i , dargestellt in Abbildung C.3.

Die Simulation der paradoxen Hitzeempfindung (Abbildung C.4) erfolgt nun durch eine einmalige Aktivierung von K_i zum Zeitpunkt $t-3$. Danach wird der Stimulus entfernt: formal wird das durch $\neg K_i(t-2)$ ausgedrückt (s. Gleichung C.1).

¹ siehe [MP43, 106]

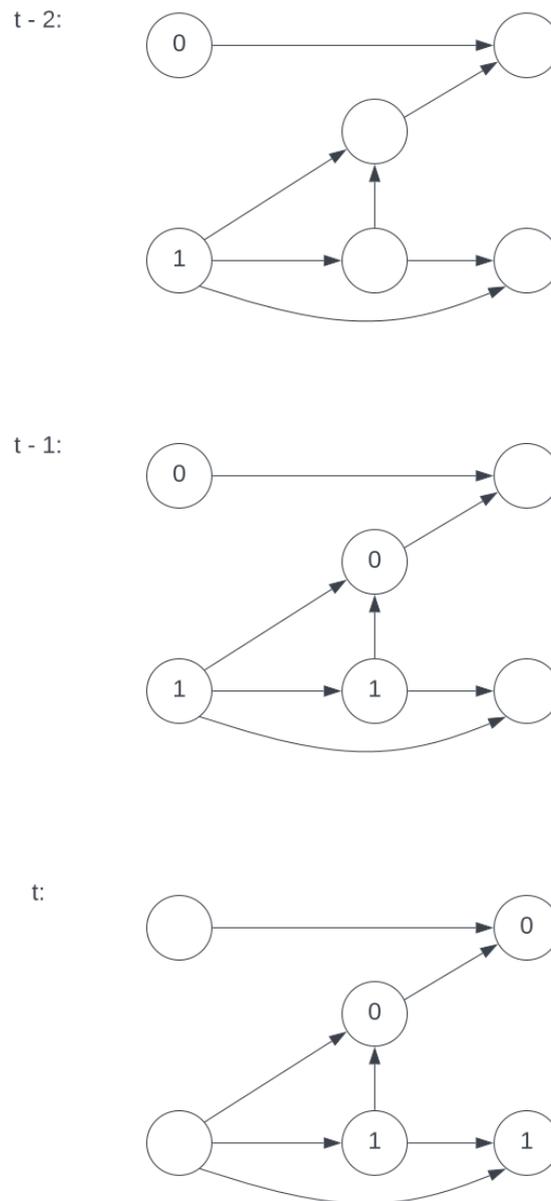


Abb. C.2. Zwei Kälteimpulse als Eingabe für das MCP-Netz. (Quelle: in Anlehnung an [Fau94, 35, Figure 1.23]).

t	H_i	K_i	Z_1	Z_2	H_o	K_o
$t - 2$	0	1				
$t - 1$	0	1	0	1		
t			0	1	0	1

Tabelle C.1. Verarbeitung von zwei Kälteimpulsen in dem MCP-Netz. Vgl. Abbildung C.2.

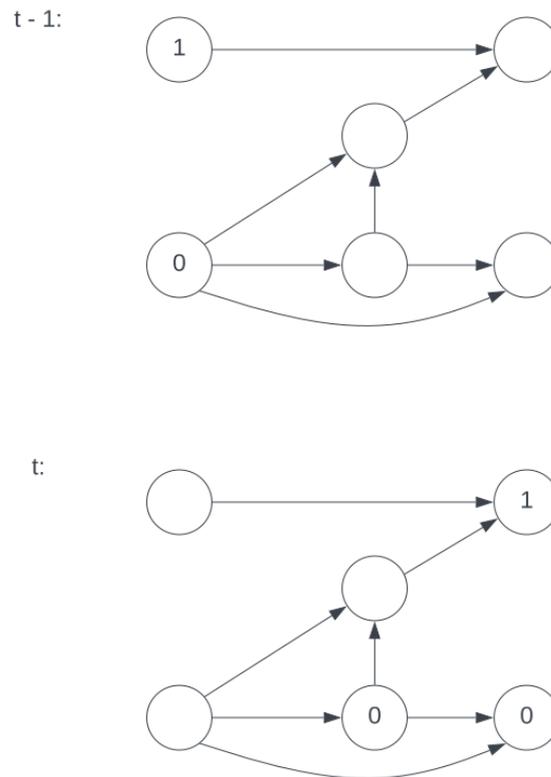


Abb. C.3. Ein einzelner Wärmeimpuls aktiviert das Wärmesignal. (Quelle: in Anlehnung an [Fau94, 36, Figure 1.24]).

t	H_i	K_i	Z_1	Z_2	H_o	K_o
$t - 1$	1	0				
t				0	1	0

Tabelle C.2. Tabellarische Darstellung von einem Wärmeimpuls. Vgl. Abbildung C.3.

t	H_i	K_i	Z_1	Z_2	H_o	K_o
$t - 3$	0	1				
$t - 2$	0	0	0	1		
$t - 1$			1	0	0	0
t					1	0

Tabelle C.3. Tabellarische Darstellung der paradoxen Hitzeempfindung. Vgl. Abbildung C.4

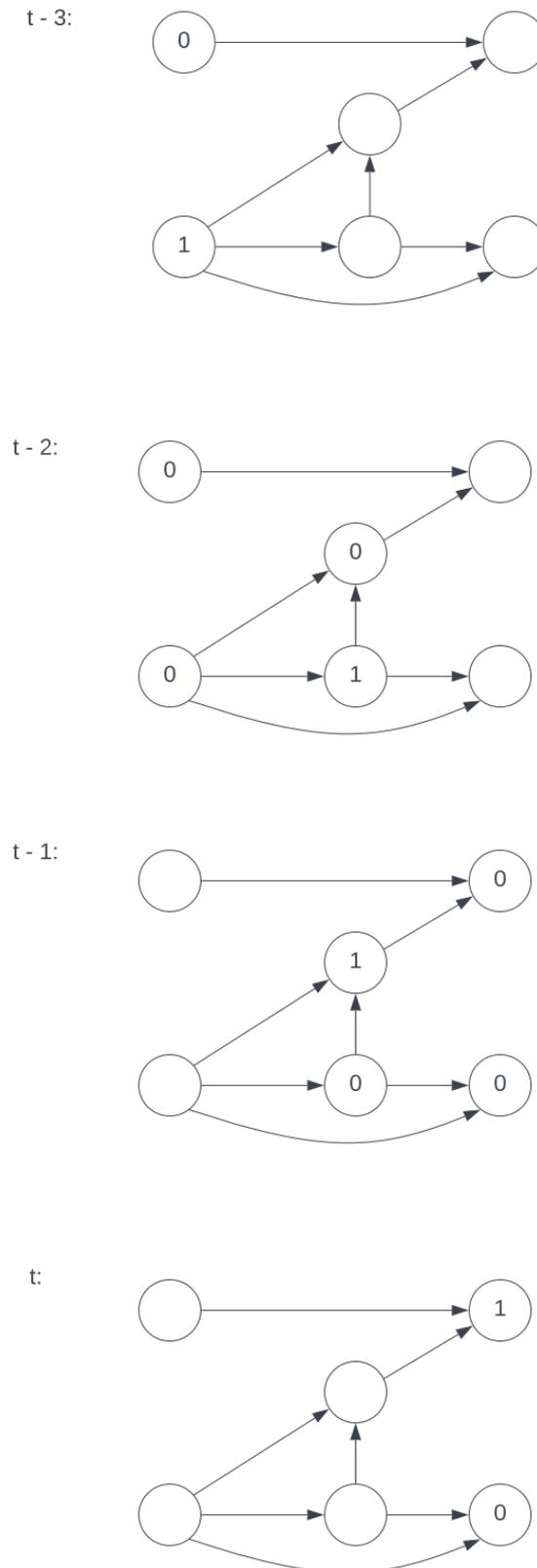


Abb. C.4. Entstehung der “paradoxen Hitzeempfindung“ durch einen einzelnen Kälteimpuls. (Quelle: in Anlehnung an [Fau94, 34, Figure 1.22]).

D

Perzeptron-Implementierung in Python

Wie bereits in Abschnitt 3.1.5 für das MCP-Neuron gezeigt, wollen wir auch mit dem Rosenblatt-Perzeptron boolesche Funktionen nachbilden.

Das folgende Beispiel bezieht sich auf die **AND**-Funktion (vgl. Tabelle 3.1).

Zunächst definieren wir zwei Mengen M_+ und M_- : Für Elemente in M_+ ist die erwartete Ausgabe 1, für Elemente in M_- ist sie 0 (vgl. Abbildung D.1):

$$M_+ := \{(1, 1)\} \quad (\text{D.1})$$

$$M_- := \{(0, 0), (0, 1), (1, 0)\} \quad (\text{D.2})$$

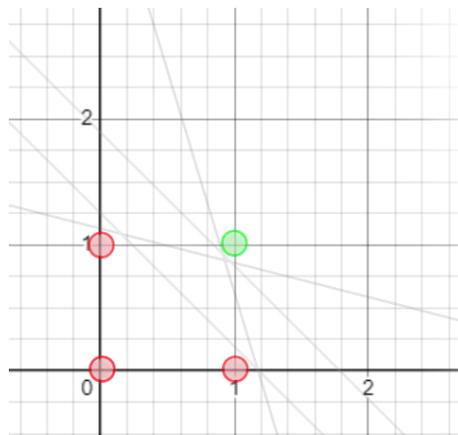


Abb. D.1. Geometrische Interpretation der Perzeptron-Funktion im \mathbb{R}^2 (Quelle: Eigene Darstellung)

Die Abbildung zeigt alle möglichen zweiwertigen Interpretationen von $A \wedge B$ in einem kartesischen Koordinatensystem. Offensichtlich ist die Anzahl der möglichen Trenngeraden (einige hier grau eingezeichnet) im \mathbb{R}^2 unendlich.

Für unsere Perzeptron-Implementierung in Python¹ implementieren wir die Lernregel wie folgt:

¹ vollständiger Quellcode unter <https://github.com/ThorstenSuckow/pylabs> (abgerufen 25.08.2023)

```

1     for epoch in range(epochs):
2
3         errors = 0
4
5         for i in range(n):
6             expected = y[i]
7             net = X[i] @ self.w + self.bias
8             result = self.heaviside(net)
9             error = 0
10
11            if result != expected:
12                error = expected - result
13
14                self.w += (X[i] * learning_rate * error)
15                self.bias += learning_rate * error

```

Der Algorithmus erhält eine vorgegebene Zahl von Epochen, in denen eine vollständige Konvergenz erwartet wird. Innerhalb einer Epoche werden alle n Trainingsdaten (enthalten in X) durchgegangen. In y finden wir die erwartete Klassifizierung der Trainingsdaten: Ist $X[i] = [1, 1]$, entspricht $y[i] = 1$ (vgl. Gleichung D.3). net entspricht dem Ergebnis von Gleichung 3.18, und das Ergebnis $result$ der **Heaviside**-Funktion mit

$$exp = \begin{cases} 1 & \text{falls } net \geq 0 \\ 0 & \text{sonst} \end{cases} \quad (\text{D.3})$$

Die Lernregel wird so lange durchlaufen, bis für alle Trainingsdaten kein Fehler $error$ mehr auftritt oder alle Epochen ausgeschöpft sind.

Der Algorithmus nutzt folgende Nomenklatur:

$\eta := 1$ (Lernrate)

$b := 0$ (Bias)

$w := \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ (Gewichtsvektor)

$M_+ := \{(1, 1)\}$

$M_- := \{(0, 0), (0, 1), (1, 0)\}$

$net = w_1x_1 + w_2x_2 + b$

$exp = \begin{cases} 1 & \text{falls } (x_1, x_2) \in M_+ \\ 0 & \text{sonst} \end{cases}$ (erwartete Ausgabe)

$res = \begin{cases} 0 & \text{falls } net < 0 \\ 1 & \text{sonst} \end{cases}$ (berechnete Ausgabe)

$err = exp - res$ (Fehler)

$$w^u = w + (x_1, x_2) * \eta * err \text{ (aktualisierter Gewichtsvektor)}$$

$$b^u = b + \eta * err \text{ (aktualisierter Bias)}$$

Mit der o.a. Implementierung konvergiert das Perzeptron innerhalb von 5 Epochen. In jeder Epoche werden 4 einzelne Daten trainiert, was insgesamt einer Anzahl von 20 Trainingsschritten entspricht. Das Perzeptron hat also für w_1, w_2 Werte gefunden, so dass die folgenden 4 Ungleichungen gleichzeitig erfüllt sind:

$$w_1 0 + w_2 0 < \Theta$$

$$w_1 1 + w_2 0 < \Theta$$

$$w_1 0 + w_2 1 < \Theta$$

$$w_1 1 + w_2 1 \geq \Theta$$

Die folgenden Tabellen zeigen die Anpassungen im Detail. Die geometrische Darstellung hierzu findet sich in Abbildung D.2.

Epoche	x_1	x_2	w_1	w_2	b	net	exp	res	err	w_1^u	w_2^u	b^u
1.1	0	0	1	1	0	0	0	1	-1	1	1	-1
1.2	1	0	1	1	-1	0	0	1	-1	0	1	-2
1.3	0	1	0	1	-2	-1	0	0	0	0	1	-2
1.4	1	1	0	1	-2	-1	1	0	1	1	2	-1

Tabelle D.1. Epoche 1 für AND

Epoche	x_1	x_2	w_1	w_2	b	net	exp	res	err	w_1^u	w_2^u	b^u
2.1	0	0	1	2	-1	-1	0	0	0	1	2	-1
2.2	1	0	1	2	-1	0	0	1	-1	0	2	-2
2.3	0	1	0	2	-2	0	0	1	-1	0	1	-3
2.4	1	1	0	1	-3	-2	1	0	1	1	2	-2

Tabelle D.2. Epoche 2 für AND

Epoche	x_1	x_2	w_1	w_2	b	net	exp	res	err	w_1^u	w_2^u	b^u
3.1	0	0	1	2	-2	-2	0	0	0	1	2	-2
3.2	1	0	1	2	-2	-1	0	0	0	1	2	-2
3.3	0	1	1	2	-2	0	0	1	-1	1	1	-3
3.4	1	1	1	1	-3	-1	1	0	1	2	2	-2

Tabelle D.3. Epoche 3 für AND

Epoche	x_1	x_2	w_1	w_2	b	net	exp	res	err	w_1^u	w_2^u	b^u
4.1	0	0	2	2	-2	-2	0	0	0	2	2	-2
4.2	1	0	2	2	-2	0	0	1	-1	1	2	-3
4.3	0	1	1	2	-3	-1	0	0	0	1	2	-3
4.4	1	1	1	2	-3	0	1	1	0	1	2	-3

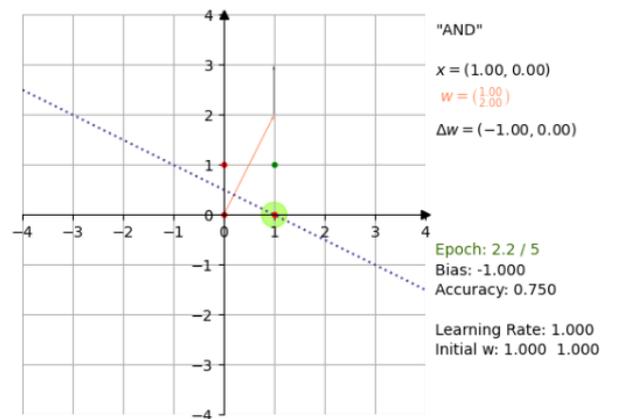
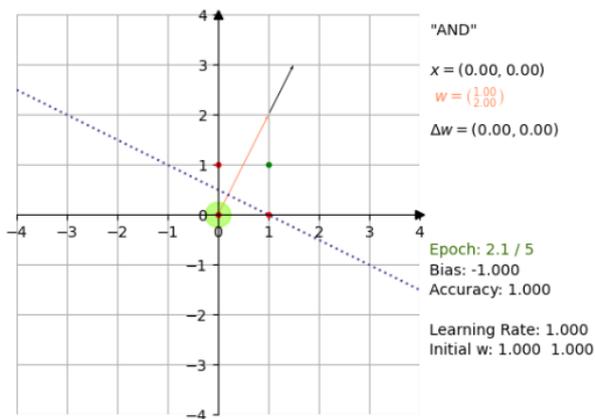
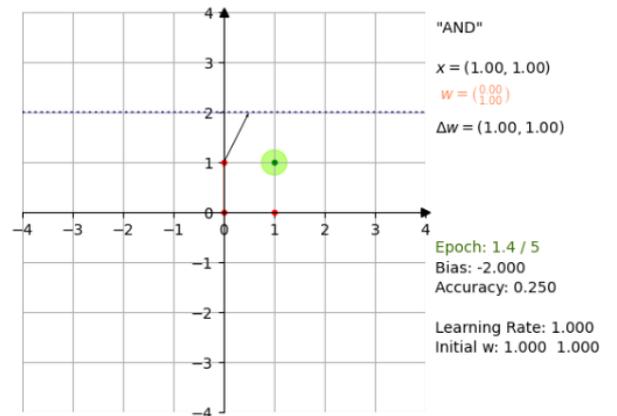
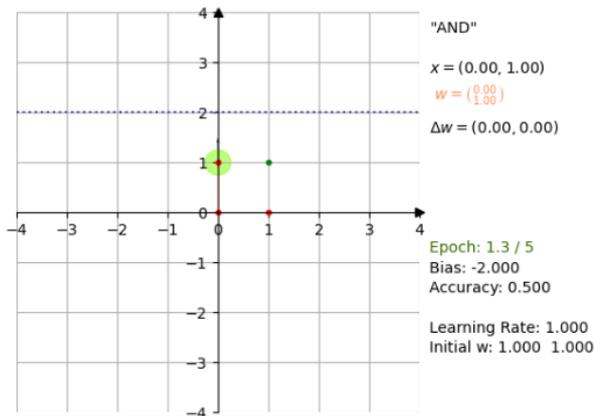
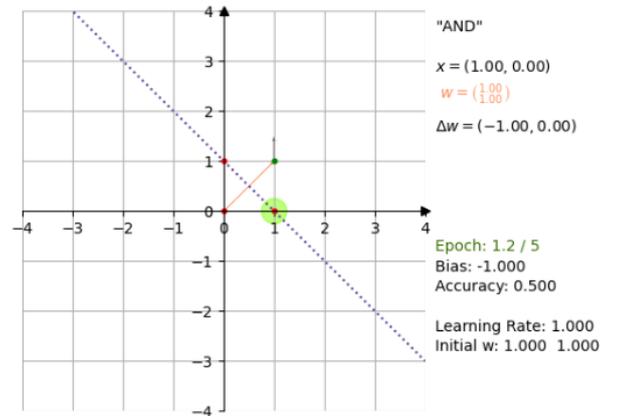
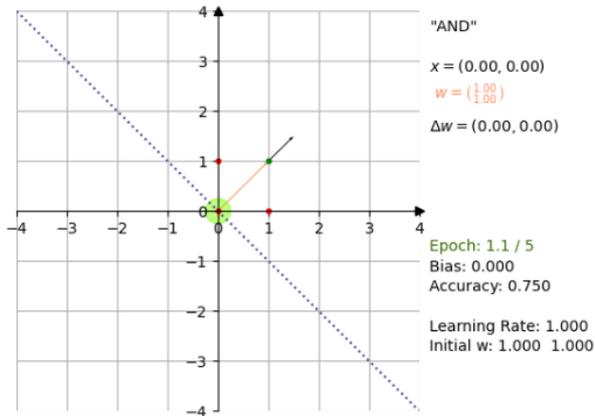
Tabelle D.4. Epoche 4 für AND

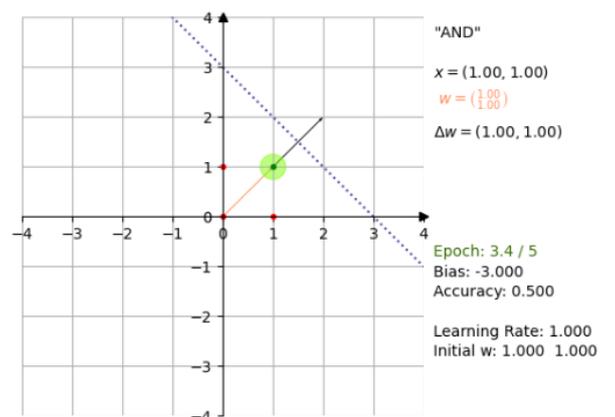
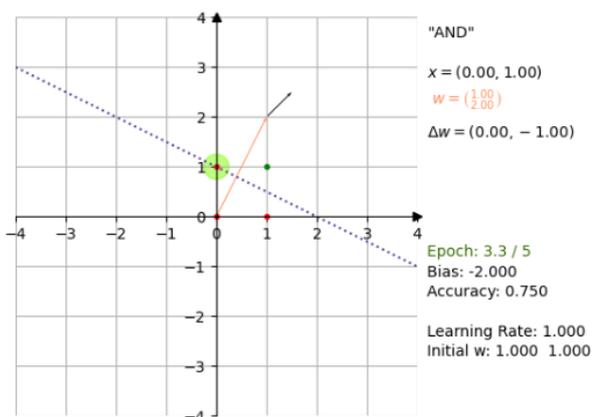
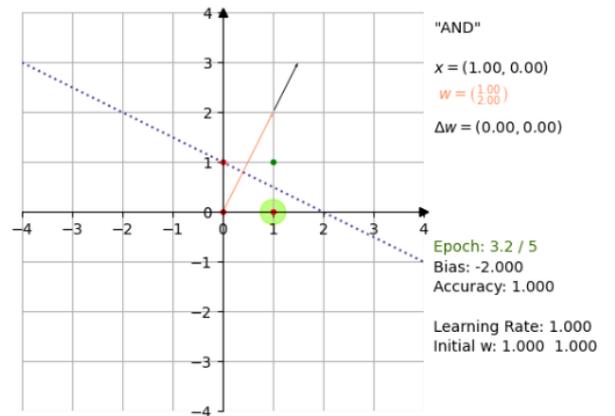
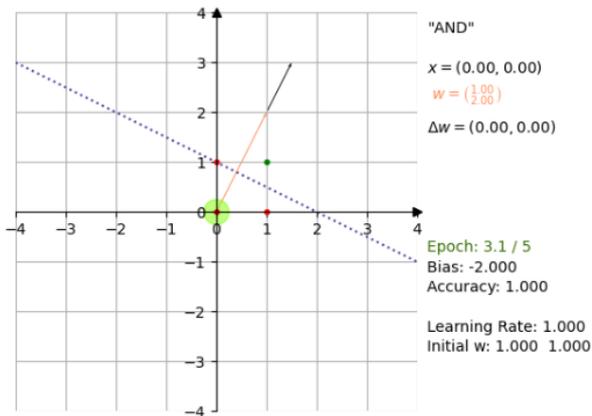
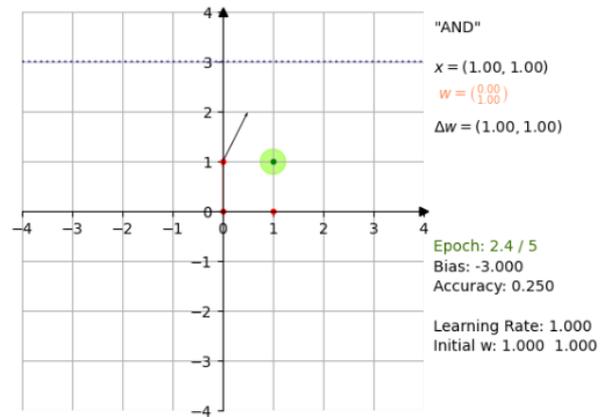
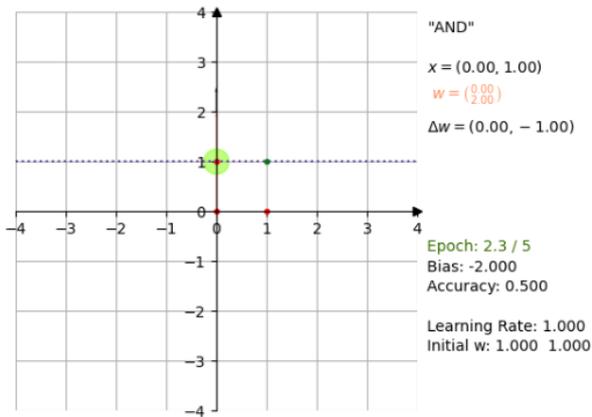
Epoche	x_1	x_2	w_1	w_2	b	net	exp	res	err	w_1^u	w_2^u	b^u
5.1	0	0	1	2	-3	-3	0	0	0	1	2	-3
5.2	1	0	1	2	-3	-2	0	0	0	1	2	-3
5.3	0	1	1	2	-3	-1	0	0	0	1	2	-3
5.4	1	1	1	2	-3	0	1	1	0	1	2	-3

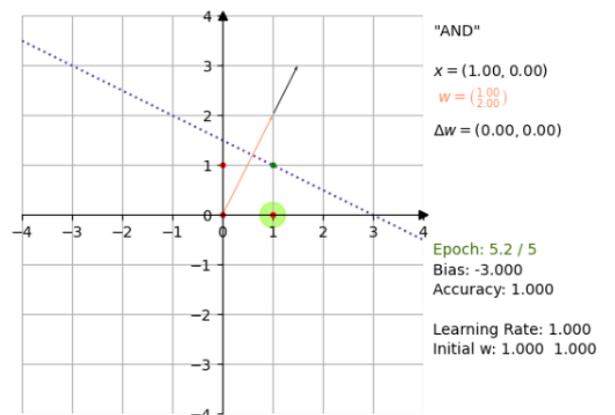
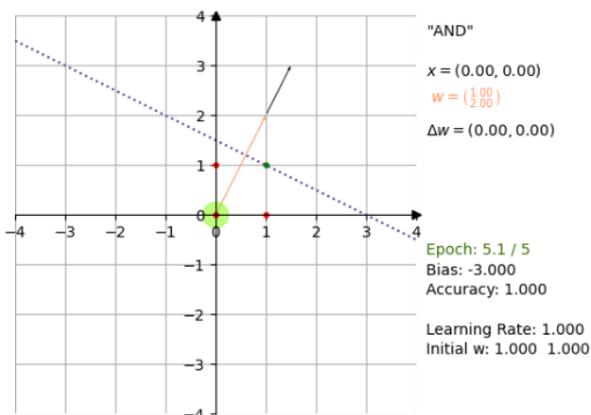
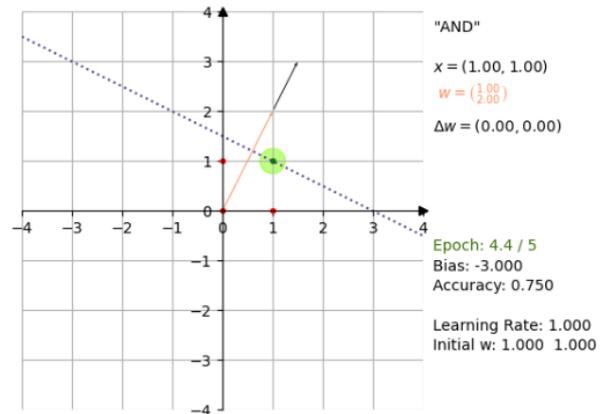
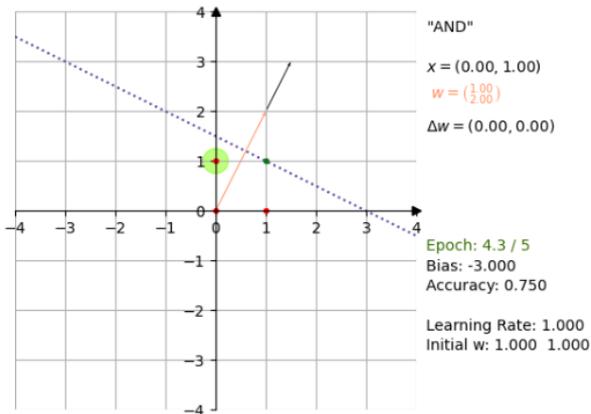
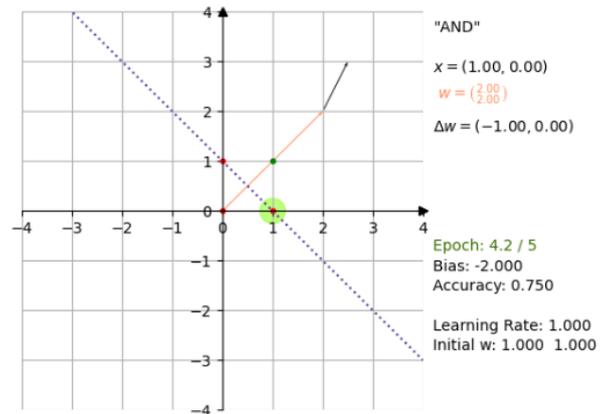
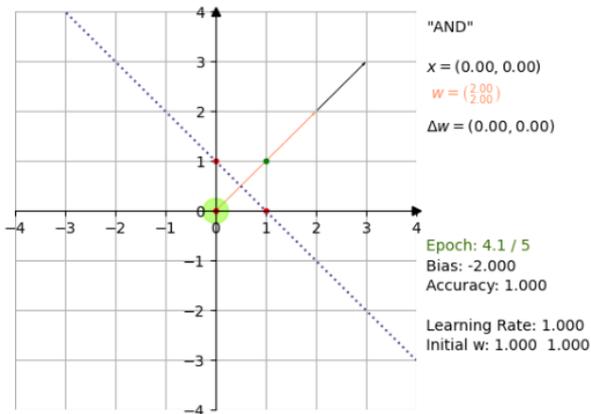
Tabelle D.5. Epoche 5 für AND

Geometrische Darstellung

Die folgende Abbildung D.2 zeigt einen vollständigen Lauf über 5 Epochen für die AND-Funktion. Blaue gestrichelt ist die Trenngerade, auf der senkrecht der Gewichtsvektor (orange) steht. Grün umrandete Punkte sind Ziel des Trainings-schrittes und bedingen den "Änderungsvektor" (schwarz), der die Anpassung von w_i im nächsten Schritt angibt.







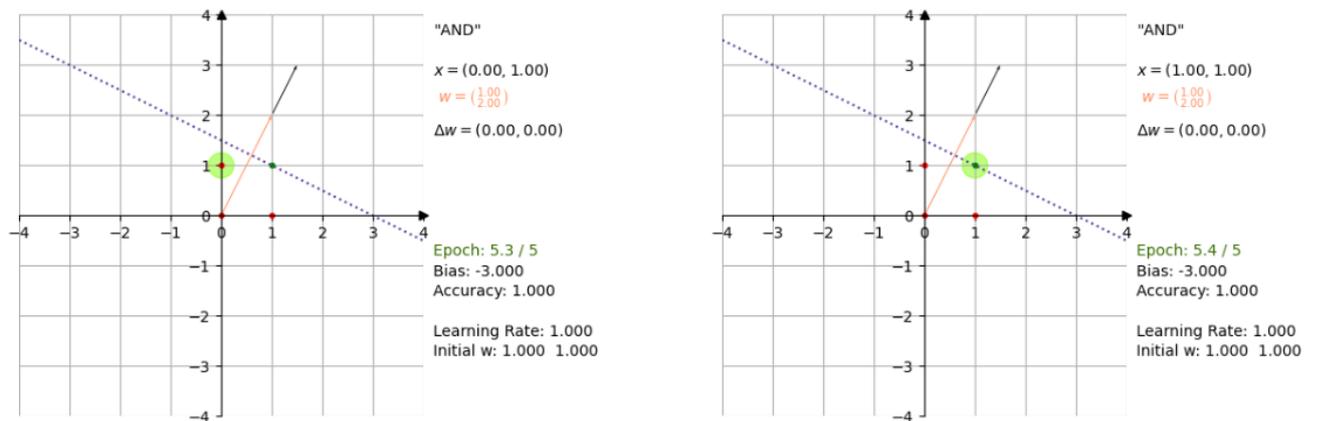


Abb. D.2. Geometrische Darstellung der Konvergenz eines Perzeptrons für AND
(Quelle: Eigene Darstellung)

Literaturverzeichnis

- [Abr02] Tara H. Abraham. „(Physio)Logical Circuits: The Intellectual Origins of the McCulloch-Pitts Neural Networks“. In: *Journal of the History of the Behavioral Sciences* 38.1 (2002), S. 3–25. ISSN: 0022-5061, 1520-6696. DOI: 10.1002/jhbs.1094.
- [AHR19] Christoph Auer, Nora Hollenstein und Matthias Reumann. „Künstliche Intelligenz im Gesundheitswesen“. In: *Gesundheit digital: Perspektiven zur Digitalisierung im Gesundheitswesen*. Hrsg. von Robin Haring. Berlin, Heidelberg: Springer Berlin Heidelberg, 2019, S. 33–46. ISBN: 978-3-662-57611-3. DOI: 10.1007/978-3-662-57611-3_3.
- [AHS85] David H. Ackley, Geoffrey E. Hinton und Terrence J. Sejnowski. „A Learning Algorithm for Boltzmann Machines“. In: *Cognitive Science* 9.1 (1985), S. 147–169. ISSN: 0364-0213. DOI: 10.1016/S0364-0213(85)80012-4.
- [Ama+13] Filippo Amato u. a. „Artificial Neural Networks in Medical Diagnosis“. In: *Journal of Applied Biomedicine* 11.2 (Juli 2013), S. 47–58. ISSN: 1214021X, 12140287. DOI: 10.2478/v10136-012-0031-x.
- [AR88] James A. Anderson und Edward Rosenfeld, Hrsg. *Neurocomputing: Foundations of Research*. Cambridge, MA, USA: MIT Press, 1988. ISBN: 0-262-01097-6.
- [AR98] James A. Anderson und Edward Rosenfeld, Hrsg. *Talking Nets: An oral history of neural networks*. Cambridge, MA, USA: MIT Press, 1998. ISBN: 0-262-01167-0.
- [Arb00] Michael A. Arbib. „Warren McCulloch’s Search for the Logic of the Nervous System“. In: *Perspectives in Biology and Medicine* 43.2 (2000), S. 193–216. ISSN: 1529-8795. DOI: 10.1353/pbm.2000.0001.
- [Arb03] Michael A. Arbib, Hrsg. *The Handbook of Brain Theory and Neural Networks*. 2nd ed. Cambridge, Mass: MIT Press, 2003. ISBN: 978-0-262-01197-6.
- [Aya+98] S. Ayad u. a. *The Extracellular Matrix. FactsBook*. 2. Aufl. London: Academic Press, 1998. ISBN: 0-12-068911-1.
- [BCP18] Mark F. Bear, Barry W. Connors und Michael A. Paradiso. *Neurowissenschaften: Ein Grundlegendes Lehrbuch Für Biologie, Medizin Und Psychologie*. Hrsg. von Andreas K. Engel. Berlin, Heidelberg: Springer Berlin Heidelberg, 2018. ISBN: 978-3-662-57263-4. DOI: 10.1007/978-3-662-57263-4.
- [Ber02] Julius Bernstein. „Untersuchungen Zur Thermodynamik Der Bioelektrischen Ströme“. In: *Archiv für die gesamte Physiologie des Menschen und der Tiere* 92.10 (Nov. 1902), S. 521–562. ISSN: 1432-2013. DOI: 10.1007/BF01790181.
- [Ber12] Julius Bernstein. „Die Membrantheorie“. In: *Elektrobiologie: Die Lehre von Den Elektrischen Vorgängen Im Organismus Auf Moderner Grundlage Dargestellt*. Wiesbaden: Vieweg+Teubner

- Verlag, 1912, S. 87–107. ISBN: 978-3-663-01627-4. DOI: 10.1007/978-3-663-01627-4_5.
- [Ber68] J. Bernstein. „Ueber Den Zeitlichen Verlauf Der Negativen Schwankung Des Nervenstroms“. In: *Archiv für die gesamte Physiologie des Menschen und der Tiere* 1.1 (Dez. 1868), S. 173–207. ISSN: 1432-2013. DOI: 10.1007/BF01640316.
- [Bis06] C. M. Bishop. *Pattern Recognition and Machine Learning*. New York: Springer, 2006.
- [BJ90] Russell Beale und Tom Jackson. *Neural Computing: An Introduction*. A Taylor & Francis Book. New York: IOP Publishing, 1990. ISBN: 0-85274-262-2.
- [BM03] Richard E. Brown und Peter M. Milner. „The Legacy of Donald O. Hebb: More than the Hebb Synapse“. In: *Nature Reviews Neuroscience* 4.12 (Dez. 2003), S. 1013–1019. ISSN: 1471-0048. DOI: 10.1038/nrn1257.
- [Bri+18] Titus Josef Brinker u. a. „Skin Cancer Classification Using Convolutional Neural Networks: Systematic Review“. In: *J Med Internet Res* 20.10 (Okt. 2018), e11936. ISSN: 1438-8871. DOI: 10.2196/11936. PMID: 30333097.
- [Bur+12] Klemens Burg u. a. „Vektorräume Beliebiger Dimensionen“. In: *Höhere Mathematik Für Ingenieure Band II: Lineare Algebra*. Wiesbaden: Vieweg+Teubner Verlag, 2012, S. 75–146. ISBN: 978-3-8348-2267-3. DOI: 10.1007/978-3-8348-2267-3_2.
- [BY86] K. Binder und A. P. Young. „Spin Glasses: Experimental Facts, Theoretical Concepts, and Open Questions“. In: *Rev. Mod. Phys.* 58.4 (Okt. 1986), S. 801–976. DOI: 10.1103/RevModPhys.58.801.
- [Cal20] Ewen Callaway. „‘It Will Change Everything’: DeepMind’s AI Makes Gigantic Leap in Solving Protein Structures“. In: *Nature* 588 (Dez. 2020), S. 203–204. DOI: 10.1038/d41586-020-03348-4.
- [CC39] Kenneth S. Cole und Howard J. Curtis. „Electric Impedance of the Squid Giant Axon during Activity“. In: *Journal of General Physiology* 22.5 (Mai 1939), S. 649–670. ISSN: 0022-1295. DOI: 10.1085/jgp.22.5.649.
- [Cho+16] Edward Choi u. a. „Doctor AI: Predicting Clinical Events via Recurrent Neural Networks“. In: *Proceedings of the 1st Machine Learning for Healthcare Conference*. Hrsg. von Finale Doshi-Velez u. a. Bd. 56. Proceedings of Machine Learning Research. Northeastern University, Boston, MA, USA: PMLR, Aug. 2016, S. 301–318. DOI: 10.48550/arXiv.1511.05942.
- [Chr73] Christian B. Anfinsen. „Principles That Govern the Folding of Protein Chains“. In: *Science (New York, N.Y.)* 181.4096 (1973), S. 223–230. DOI: 10.1126/science.181.4096.223.
- [CL78] Martha Constantine-Paton und Margaret I. Law. „Eye-Specific Termination Bands in Tecta of Three-Eyed Frogs“. In: *Science* 202.4368 (1978), S. 639–641. DOI: 10.1126/science.309179.

- [Coo05] Steven J. Cooper. „Donald O. Hebb’s Synapse and Learning Rule: A History and Commentary“. In: *Neuroscience & Biobehavioral Reviews* 28.8 (2005), S. 851–874. ISSN: 0149-7634. DOI: 10.1016/j.neubiorev.2004.09.009.
- [Cow90] Jack D. Cowan. „Discussion: McCulloch-Pitts and Related Neural Nets from 1943 to 1989“. In: *Bulletin of Mathematical Biology* 52.1 (Jan. 1990), S. 73–97. ISSN: 1522-9602. DOI: 10.1007/BF02459569.
- [Das+19] Sabyasachi Dash u. a. „Big Data in Healthcare: Management, Analysis and Future Prospects“. In: *Journal of Big Data* 6.1 (Juni 2019), S. 54. ISSN: 2196-1115. DOI: 10.1186/s40537-019-0217-0.
- [Die17] R. Diestel. *Graphentheorie*. Heidelberg: Springer-Verlag, 2017. ISBN: 9783961340040.
- [Dil+08] Ken A. Dill u. a. „The Protein Folding Problem“. In: *Annual Review of Biophysics* 37.1 (2008), S. 289–316. DOI: 10.1146/annurev.biophys.37.092707.153558.
- [Dor91] Georg Dorffner. „Konnektionismus — Eine Einführung“. In: *Konnektionismus: Von Neuronalen Netzwerken Zu Einer „natürlichen“ KI*. Wiesbaden: Vieweg+Teubner Verlag, 1991, S. 15–83. ISBN: 978-3-322-94665-2. DOI: 10.1007/978-3-322-94665-2_2.
- [Eil19] J. Eilers. „Nervenzellen“. In: *Physiologie Des Menschen: Mit Pathophysiologie*. Hrsg. von Ralf Brandes, Florian Lang und Robert F. Schmidt. Berlin, Heidelberg: Springer Berlin Heidelberg, 2019, S. 57–64. ISBN: 978-3-662-56468-4. DOI: 10.1007/978-3-662-56468-4_5.
- [Ert21a] Wolfgang Ertel. „Lernen Durch Verstärkung (Reinforcement Learning)“. In: *Grundkurs Künstliche Intelligenz: Eine Praxisorientierte Einführung*. Wiesbaden: Springer Fachmedien Wiesbaden, 2021, S. 351–377. ISBN: 978-3-658-32075-1. DOI: 10.1007/978-3-658-32075-1_10.
- [Ert21b] Wolfgang Ertel. „Maschinelles Lernen Und Data Mining“. In: *Grundkurs Künstliche Intelligenz: Eine Praxisorientierte Einführung*. Wiesbaden: Springer Fachmedien Wiesbaden, 2021, S. 201–283. ISBN: 978-3-658-32075-1. DOI: 10.1007/978-3-658-32075-1_8.
- [Ert21c] Wolfgang Ertel. „Neuronale Netze“. In: *Grundkurs Künstliche Intelligenz: Eine Praxisorientierte Einführung*. Wiesbaden: Springer Fachmedien Wiesbaden, 2021, S. 285–349. ISBN: 978-3-658-32075-1. DOI: 10.1007/978-3-658-32075-1_9.
- [Est+17] Andre Esteva u. a. „Dermatologist-Level Classification of Skin Cancer with Deep Neural Networks“. In: *Nature* 542.7639 (Feb. 2017), S. 115–118. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/nature21056.
- [Fak19] B. Fakler. „Grundlagen Der Zellulären Erregbarkeit“. In: *Physiologie Des Menschen: Mit Pathophysiologie*. Hrsg. von Ralf Brandes, Florian Lang und Robert F. Schmidt. Berlin, Heidelberg: Springer Berlin Heidelberg, 2019, S. 38–54. ISBN: 978-3-662-56468-4. DOI: 10.1007/978-3-662-56468-4_4.

- [Fau94] Laurene Fausett. *Fundamentals of Neural Networks: Architectures, Algorithms, and Applications*. USA: Prentice-Hall, Inc., 1994. ISBN: 0-13-334186-0.
- [Fay+19] O.S.I. Fayomi u. a. „Effect of Annealing on the Mechanical Characteristics of Steel Welded Joint“. In: *Procedia Manufacturing* 35 (2019), S. 1387–1394. ISSN: 23519789. DOI: 10.1016/j.promfg.2019.09.008.
- [FC54] B. Farley und W. Clark. „Simulation of Self-Organizing Systems by Digital Computer“. In: *Transactions of the IRE Professional Group on Information Theory* 4.4 (1954), S. 76–84. DOI: 10.1109/TIT.1954.1057468.
- [FE19] B. Fakler und J. Eilers. „Ruhemembranpotenzial Und Aktionspotenzial“. In: *Physiologie Des Menschen: Mit Pathophysiologie*. Hrsg. von Ralf Brandes, Florian Lang und Robert F. Schmidt. Berlin, Heidelberg: Springer Berlin Heidelberg, 2019, S. 65–71. ISBN: 978-3-662-56468-4. DOI: 10.1007/978-3-662-56468-4_6.
- [Flo19] H. Flor. „Lernen“. In: *Physiologie Des Menschen: Mit Pathophysiologie*. Hrsg. von Ralf Brandes, Florian Lang und Robert F. Schmidt. Berlin, Heidelberg: Springer Berlin Heidelberg, 2019, S. 827–838. ISBN: 978-3-662-56468-4. DOI: 10.1007/978-3-662-56468-4_66.
- [FMI83] Kunihiko Fukushima, Sei Miyake und Takayuki Ito. „Neocognitron: A Neural Network Model for a Mechanism of Visual Pattern Recognition“. In: *IEEE Transactions on Systems, Man, and Cybernetics* SMC-13.5 (1983), S. 826–834. DOI: 10.1109/TSMC.1983.6313076.
- [Fra94] Robert G. Frank. „Instruments, Nerve Action, and the All-or-None Principle“. In: *Osiris* 9.1 (1994), S. 208–235. DOI: 10.1086/368737.
- [Fro19] M. Fromm. „Transport in Membranen Und Epithelien“. In: *Physiologie Des Menschen: Mit Pathophysiologie*. Hrsg. von Ralf Brandes, Florian Lang und Robert F. Schmidt. Berlin, Heidelberg: Springer Berlin Heidelberg, 2019, S. 22–37. ISBN: 978-3-662-56468-4. DOI: 10.1007/978-3-662-56468-4_3.
- [FS90] Gerhard Friedrich und Markus Stumptner. „Einführung“. In: *Expertensysteme*. Hrsg. von Georg Gottlob, Thomas Frühwirth und Werner Horn. Vienna: Springer Vienna, 1990, S. 1–19. ISBN: 978-3-7091-9094-4. DOI: 10.1007/978-3-7091-9094-4_1.
- [Fuk75] Kunihiko Fukushima. „Cognitron: A Self-Organizing Multilayered Neural Network“. In: *Biological Cybernetics* 20.3 (Sep. 1975), S. 121–136. ISSN: 1432-0770. DOI: 10.1007/BF00342633.
- [Fuk80] Kunihiko Fukushima. „Neocognitron: A Self-Organizing Neural Network Model for a Mechanism of Pattern Recognition Unaffected by Shift in Position“. In: *Biological Cybernetics* 36.4 (Apr. 1980), S. 193–202. ISSN: 1432-0770. DOI: 10.1007/BF00344251.
- [Gar19] Colin Garvey. „Artificial Intelligence and Japan’s Fifth Generation: The Information Society, Neoliberalism, and Alternative Modernities“. In: *Pacific Historical Review* 88.4 (Nov. 2019), S. 619–658. ISSN: 0030-8684. DOI: 10.1525/phr.2019.88.4.619.

- [GBC18] Ian Goodfellow, Yoshua Bengio und Aaron Courville. *Deep Learning. Das umfassende Handbuch*. 2018. Aufl. mitp Professional. Frechen, Germany: MITP, Okt. 2018. ISBN: 978-3-95845-700-3.
- [Gle09] P. Glees. „Nobelpreis für Physiologie und Medizin 1963“. In: *DMW - Deutsche Medizinische Wochenschrift* 88.52 (Apr. 2009), S. 2523–2526. ISSN: 0012-0472. DOI: 10.1055/s-0028-1112385.
- [Häm+82] H. Härmäläinen u. a. „Paradoxical Heat Sensations during Moderate Cooling of the Skin“. In: *Brain Research* 251.1 (1982), S. 77–81. ISSN: 0006-8993. DOI: 10.1016/0006-8993(82)91275-6.
- [Heb49] D. O. Hebb. *The Organization of Behavior*. New York: Wiley, 1949.
- [Heb88] Donald O. Hebb. „The Organization of Behavior“. In: *Neurocomputing: Foundations of Research*. Cambridge, MA, USA: MIT Press, 1988, S. 43–54. ISBN: 0-262-01097-6.
- [HH52] A. L. Hodgkin und A. F. Huxley. „A Quantitative Description of Membrane Current and Its Application to Conduction and Excitation in Nerve“. In: *The Journal of Physiology* 117.4 (1952), S. 500–544. DOI: 10.1113/jphysiol.1952.sp004764.
- [HI97] Frank C. Hoppensteadt und Eugene M. Izhikevich. „Introduction (Aus: Weakly Connected Neural Networks)“. In: *Weakly Connected Neural Networks*. New York, NY: Springer New York, 1997, S. 3–24. ISBN: 978-1-4612-1828-9. DOI: 10.1007/978-1-4612-1828-9_1.
- [HLW16] Gao Huang, Zhuang Liu und Kilian Q. Weinberger. „Densely Connected Convolutional Networks“. In: *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2016), S. 2261–2269.
- [Hof22] Dirk W. Hoffmann. *Theoretische Informatik*. 5., aktualisierte Auflage. München: Hanser, 2022. ISBN: 978-3-446-47029-3.
- [Hop82] John Hopfield. „Neural Networks and Physical Systems with Emergent Collective Computational Abilities“. In: *Proceedings of the National Academy of Sciences of the United States of America* 79 (Mai 1982), S. 2554–8. DOI: 10.1073/pnas.79.8.2554.
- [Hor90] Werner Horn. „Anwendungen von Expertensystemen“. In: *Expertensysteme*. Hrsg. von Georg Gottlob, Thomas Frühwirth und Werner Horn. Vienna: Springer Vienna, 1990, S. 61–72. ISBN: 978-3-7091-9094-4. DOI: 10.1007/978-3-7091-9094-4_3.
- [HS19a] S. Hallermann und R. F. Schmidt. „Arbeitsweise von Synapsen“. In: *Physiologie Des Menschen: Mit Pathophysiologie*. Hrsg. von Ralf Brandes, Florian Lang und Robert F. Schmidt. Berlin, Heidelberg: Springer Berlin Heidelberg, 2019, S. 95–104. ISBN: 978-3-662-56468-4. DOI: 10.1007/978-3-662-56468-4_9.
- [HS19b] S. Hallermann und R. F. Schmidt. „Neurotransmitter Und Ihre Rezeptoren“. In: *Physiologie Des Menschen: Mit Pathophysiologie*. Hrsg. von Ralf Brandes, Florian Lang und Robert F. Schmidt. Berlin, Heidelberg: Springer Berlin Heidelberg, 2019, S. 105–114. ISBN: 978-3-662-56468-4. DOI: 10.1007/978-3-662-56468-4_10.

- [HS19c] S. Hallermann und R. F. Schmidt. „Synaptische Plastizität“. In: *Physiologie Des Menschen: Mit Pathophysiologie*. Hrsg. von Ralf Brandes, Florian Lang und Robert F. Schmidt. Berlin, Heidelberg: Springer Berlin Heidelberg, 2019, S. 115–120. ISBN: 978-3-662-56468-4. DOI: 10.1007/978-3-662-56468-4_11.
- [HS83] Geoffrey Hinton und Terrence Sejnowski. „Optimal Perceptual Inference“. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Jan. 1983, S. 448–453.
- [HW62] D. H. Hubel und T. N. Wiesel. „Receptive Fields, Binocular Interaction and Functional Architecture in the Cat’s Visual Cortex“. In: *The Journal of Physiology* 160.1 (Jan. 1962), S. 106–154. ISSN: 00223751. DOI: 10.1113/jphysiol.1962.sp006837.
- [HWN22] Tobias Höfer, Frederik Weishäupl und Alfred Nischwitz. „Neuronale Netze Zur Effizienzsteigerung Der Texterkennung in Der Rezeptabrechnung“. In: *Künstliche Intelligenz Im Gesundheitswesen: Entwicklungen, Beispiele Und Perspektiven*. Hrsg. von Mario A. Pfannstiel. Wiesbaden: Springer Fachmedien Wiesbaden, 2022, S. 697–714. ISBN: 978-3-658-33597-7. DOI: 10.1007/978-3-658-33597-7_33.
- [Irv+19] Jeremy Irvin u. a. *CheXpert: A Large Chest Radiograph Dataset with Uncertainty Labels and Expert Comparison*. Jan. 2019. arXiv: 1901.07031 [cs, eess].
- [Jon19] P. Jonas. „Aktionspotenzial: Fortleitung Im Axon“. In: *Physiologie Des Menschen: Mit Pathophysiologie*. Hrsg. von Ralf Brandes, Florian Lang und Robert F. Schmidt. Berlin, Heidelberg: Springer Berlin Heidelberg, 2019, S. 72–82. ISBN: 978-3-662-56468-4. DOI: 10.1007/978-3-662-56468-4_7.
- [Jum+21] John Jumper u. a. „Highly Accurate Protein Structure Prediction with AlphaFold“. In: *Nature* 596.7873 (Aug. 2021), S. 583–589. ISSN: 1476-4687. DOI: 10.1038/s41586-021-03819-2.
- [Kan+13] E.R. Kandel u.a. *Principles of Neural Science, 5th Edition*. McGraw-Hill’s AccessMedicine. McGraw-Hill Education, 2013. ISBN: 978-0-07-139011-8.
- [KB17] Diederik P. Kingma und Jimmy Ba. *Adam: A Method for Stochastic Optimization*. Jan. 2017. arXiv: 1412.6980 [cs].
- [KG14] Christian Keysers und Valeria Gazzola. „Hebbian Learning and Predictive Mirror Neurons for Actions, Sensations and Emotions“. In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 369.1644 (2014), S. 20130175. DOI: 10.1098/rstb.2013.0175.
- [Kle99] Raymond M. Klein. „The Hebb Legacy.“ In: *Canadian Journal of Experimental Psychology / Revue canadienne de psychologie expérimentale* 53.1 (1999), S. 1–3. ISSN: 1878-7290(Electronic),1196-1961(Print). DOI: 10.1037/h0087295.
- [Koc98] Christof Koch. „The Hodgkin-Huxley Model of Action Potential Generation“. In: *Biophysics of Computation: Information Proces-*

- sing in Single Neurons*. Oxford University Press, Nov. 1998. ISBN: 978-0-19-510491-2. DOI: 10.1093/oso/9780195139853.003.0012.
- [Köh90] Monika Köhle. *Neurale Netze*. Vienna: Springer Vienna, 1990. ISBN: 978-3-7091-9093-7.
- [KSH12] Alex Krizhevsky, Ilya Sutskever und Geoffrey E Hinton. „ImageNet Classification with Deep Convolutional Neural Networks“. In: *Advances in Neural Information Processing Systems*. Hrsg. von F. Pereira u. a. Bd. 25. Curran Associates, Inc., 2012.
- [Kub30] Lawrence S. Kubie. „A theoretical application to some neurological problems of the properties of excitation waves which move in closed circuits“. In: *Brain* 53.2 (Juli 1930), S. 166–177. ISSN: 0006-8950. DOI: 10.1093/brain/53.2.166.
- [Küp19] E. W. Udo Küppers. „Einleitung Und Lernziele“. In: *Eine Transdisziplinäre Einführung in Die Welt Der Kybernetik: Grundlagen, Modelle, Theorien Und Praxisbeispiele*. Wiesbaden: Springer Fachmedien Wiesbaden, 2019, S. 1–3. ISBN: 978-3-658-23725-7. DOI: 10.1007/978-3-658-23725-7_1.
- [Lan09] Anders Lansner. „Associative Memory Models: From the Cell-Assembly Theory to Biophysically Detailed Cortex Simulations“. In: *Trends in Neurosciences* 32.3 (2009), S. 178–186. ISSN: 0166-2236. DOI: 10.1016/j.tins.2008.12.002.
- [LBH15] Yann LeCun, Yoshua Bengio und Geoffrey Hinton. „Deep Learning“. In: *Nature* 521.7553 (Mai 2015), S. 436–444. ISSN: 1476-4687. DOI: 10.1038/nature14539.
- [LeC+89] Y. LeCun u. a. „Backpropagation Applied to Handwritten Zip Code Recognition“. In: *Neural Computation* 1.4 (Dez. 1989), S. 541–551. ISSN: 0899-7667. DOI: 10.1162/neco.1989.1.4.541.
- [LeC+98] Y. LeCun u. a. „Gradient-Based Learning Applied to Document Recognition“. In: *Proceedings of the IEEE* 86.11 (1998), S. 2278–2324. DOI: 10.1109/5.726791.
- [LeC89] Y. LeCun. „Generalization and Network Design Strategies“. In: *Connectionism in Perspective*. Hrsg. von R. Pfeifer u. a. Zurich, Switzerland: Elsevier, 1989.
- [Li+21] Rui Li u. a. „G-Net: A Recurrent Network Approach to G-Computation for Counterfactual Prediction Under a Dynamic Treatment Regime“. In: *Proceedings of Machine Learning for Health*. Hrsg. von Subhrajit Roy u. a. Bd. 158. Proceedings of Machine Learning Research. PMLR, Dez. 2021, S. 282–299.
- [Lip87] R. Lippmann. „An Introduction to Computing with Neural Nets“. In: *IEEE ASSP Magazine* 4.2 (1987), S. 4–22. DOI: 10.1109/MASSP.1987.1165576.
- [Liu10] Lydia H. Liu. „The Cybernetic Unconscious: Rethinking Lacan, Poe, and French Theory“. In: *Critical Inquiry* 36.2 (2010), S. 288–320. DOI: 10.1086/648527.

- [LS92] Siegrid Löwel und Wolf Singer. „Selection of Intrinsic Horizontal Connections in the Visual Cortex by Correlated Neuronal Activity“. In: *Science* 255.5041 (1992), S. 209–212. DOI: 10.1126/science.1372754.
- [Luc+22] Adriano Lucieri u.a. „Erklärbare KI in Der Medizinischen Diagnose – Erfolge Und Herausforderungen“. In: *Künstliche Intelligenz Im Gesundheitswesen: Entwicklungen, Beispiele Und Perspektiven*. Hrsg. von Mario A. Pfannstiel. Wiesbaden: Springer Fachmedien Wiesbaden, 2022, S. 727–754. ISBN: 978-3-658-33597-7. DOI: 10.1007/978-3-658-33597-7_35.
- [Luc05] Keith Lucas. „On the Gradation of Activity in a Skeletal Muscle-Fibre“. In: *The Journal of Physiology* 33.2 (1905), S. 125–137. DOI: 10.1113/jphysiol.1905.sp001115.
- [Mas90] P. R. Masani. „McCulloch, Pitts and the Evolution of Wiener’s Neurophysiological Ideas“. In: *Norbert Wiener 1894–1964*. Basel: Birkhäuser Basel, 1990, S. 218–238. ISBN: 978-3-0348-9252-0. DOI: 10.1007/978-3-0348-9252-0_16.
- [McC04] Pamela McCorduck. *Machines Who Think: A Personal Inquiry into the History and Prospects of Artificial Intelligence*. AK Peters Ltd, 2004. ISBN: 1-56881-205-1.
- [McC13] David A. McCormick. „Chapter 5 - Membrane Potential and Action Potential“. In: *Fundamental Neuroscience (Fourth Edition)*. Hrsg. von Larry R. Squire u.a. Fourth Edition. San Diego: Academic Press, 2013, S. 93–116. ISBN: 978-0-12-385870-2. DOI: 10.1016/B978-0-12-385870-2.00005-6.
- [McC16] Warren S McCulloch. *Embodiments of Mind*. London, England: MIT Press, Okt. 2016. ISBN: 9780262340960.
- [McD80] John P. McDermott. „RI: an Expert in the Computer Systems Domain“. In: *AAAI Conference on Artificial Intelligence*. 1980.
- [Min67] Marvin Minsky. *Computation: Finite and Infinite Machines*. USA: Prentice Hal, 1967.
- [MP43] Warren S. McCulloch und Walter Pitts. „A Logical Calculus of the Ideas Immanent in Nervous Activity“. In: *The bulletin of mathematical biophysics* 5.4 (Dez. 1943), S. 115–133. ISSN: 1522-9602. DOI: 10.1007/BF02478259.
- [MP88] Marvin Minsky und Seymour A Papert. *Perceptrons: An Introduction to Computational Geometry*. London, England: MIT Press, 1988.
- [Mur91] Fionn Murtagh. „Multilayer Perceptrons for Classification and Regression“. In: *Neurocomputing* 2.5 (1991), S. 183–197. ISSN: 0925-2312. DOI: 10.1016/0925-2312(91)90023-5.
- [Nov62] A. B. Novikoff. „On Convergence Proofs on Perceptrons“. In: *Proceedings of the Symposium on the Mathematical Theory of Automata*. Bd. 12. New York, NY, USA: Polytechnic Institute of Brooklyn, 1962, S. 615–622.
- [NP16] Hoang Nguyen und Jon Patrick. „Text Mining in Clinical Domain: Dealing with Noise“. In: *Proceedings of the 22nd ACM SIGKDD In-*

- ternational Conference on Knowledge Discovery and Data Mining. KDD '16. San Francisco, California, USA: Association for Computing Machinery, 2016, S. 549–558. ISBN: 9781450342322. DOI: 10.1145/2939672.2939720.*
- [Ola96] Mikel Olazaran. „A Sociological Study of the Official History of the Perceptrons Controversy“. In: *Social Studies of Science* 26.3 (1996), S. 611–659. DOI: 10.1177/030631296026003005.
- [OV00] Walter Oberschelp und Gottfried Vossen. *Rechneraufbau Und Rechnerstrukturen*. 8. Aufl. Oldenbourg Wissenschaftsverlag, 2000. ISBN: 3-486-25340-9.
- [Pfa22] Mario A. Pfannstiel. „Einleitung „Künstliche Intelligenz Im Gesundheitswesen““. In: *Künstliche Intelligenz Im Gesundheitswesen: Entwicklungen, Beispiele Und Perspektiven*. Hrsg. von Mario A. Pfannstiel. Wiesbaden: Springer Fachmedien Wiesbaden, 2022, S. 1–47. ISBN: 978-3-658-33597-7. DOI: 10.1007/978-3-658-33597-7_1.
- [Pic04] Gualtiero Piccinini. „The First Computational Theory of Mind and Brain: A Close Look at McCulloch and Pitts’s “Logical Calculus of Ideas Immanent in Nervous Activity”“. In: *Synthese* 141.2 (Aug. 2004), S. 175–215. ISSN: 1573-0964. DOI: 10.1023/B:SYNT.0000043018.52445.3e.
- [Raj+18] Alvin Rajkomar u. a. „Scalable and Accurate Deep Learning with Electronic Health Records“. In: *npj Digital Medicine* 1.1 (Mai 2018), S. 18. ISSN: 2398-6352. DOI: 10.1038/s41746-018-0029-1.
- [Rau08] Wolfgang Rautenberg. „Aussagenlogik“. In: *Einführung in Die Mathematische Logik: Ein Lehrbuch*. Wiesbaden: Vieweg+Teubner, 2008, S. 1–32. ISBN: 978-3-8348-9530-1. DOI: 10.1007/978-3-8348-9530-1_1.
- [RHW86] David E. Rumelhart, Geoffrey E. Hinton und Ronald J. Williams. „Learning Representations by Back-Propagating Errors“. In: *Nature* 323.6088 (Okt. 1986), S. 533–536. ISSN: 1476-4687. DOI: 10.1038/323533a0.
- [RK18] Reinhard Rohkamm und Pawel Kermer. *Taschenatlas Neurologie*. Georg Thieme Verlag, 2018. DOI: 10.1055/b-005-143299.
- [RM87] David E. Rumelhart und James L. McClelland. „PDP Models and General Issues in Cognitive Science“. In: *Parallel Distributed Processing: Explorations in the Microstructure of Cognition: Foundations*. 1987, S. 110–146.
- [RN12] Stuart Russell und Peter Norvig. *Künstliche Intelligenz*. 3. Aufl. Pearson Studium - IT. Munich, Germany: Pearson Studium ein Imprint von Pearson Deutschland, Juni 2012. ISBN: 978-3-86894-098-5.
- [Roj93] Raúl Rojas. *Theorie der neuronalen Netze*. Springer Berlin Heidelberg, 1993. DOI: 10.1007/978-3-642-61231-2.
- [Ros57] F. Rosenblatt. *The Perceptron - A Perceiving and Recognizing Automaton*. Techn. Ber. 85-460-1. Ithaca, New York: Cornell Aeronautical Laboratory, Jan. 1957.

- [Ros58] F. Rosenblatt. „The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain.“ In: *Psychological Review* 65.6 (1958), S. 386–408. ISSN: 1939-1471(Electronic),0033-295X(Print). DOI: 10.1037/h0042519.
- [Ros60] Frank Rosenblatt. „Perceptron Simulation Experiments“. In: *Proceedings of the IRE* 48.3 (1960), S. 301–309. DOI: 10.1109/JRPROC.1960.287598.
- [Ros62] Frank Rosenblatt. *Principles of neurodynamics: Perceptrons and the theory of brain mechanisms*. Bd. 55. Spartan books Washington, DC, 1962.
- [RS19] Stefan Rüping und Jil Sander. „Big Data in Gesundheitswesen und Medizin“. In: *Gesundheit digital: Perspektiven zur Digitalisierung im Gesundheitswesen*. Hrsg. von Robin Haring. Berlin, Heidelberg: Springer Berlin Heidelberg, 2019, S. 15–31. ISBN: 978-3-662-57611-3. DOI: 10.1007/978-3-662-57611-3_2.
- [Sal90] R. Salomon. „Beschleunigtes Lernen Durch Adaptive Regelung Der Lernrate Bei Back-Propagation in Feed-Forward Netzen“. In: *Konnektionismus in Artificial Intelligence Und Kognitionsforschung*. Hrsg. von Georg Dorffner. Berlin, Heidelberg: Springer Berlin Heidelberg, 1990, S. 173–178. ISBN: 978-3-642-76070-9.
- [Sch83] Stephen M. Schuetze. „The discovery of the action potential“. In: *Trends in Neurosciences* 6 (1983), S. 164–168. DOI: 10.1016/0166-2236(83)90078-4.
- [SD07] Stefan Silbernagl und Agamemnon Despopoulos, Hrsg. *Taschenatlas Physiologie*. Georg Thieme Verlag, 2007. DOI: 10.1055/b-002-89576.
- [Sey06] Ernst-August Seyfarth. „Julius Bernstein (1839–1917): Pioneer Neurobiologist and Biophysicist“. In: *Biological Cybernetics* 94.1 (Jan. 2006), S. 2–8. ISSN: 1432-0770. DOI: 10.1007/s00422-005-0031-y.
- [Sha+20] Zach Shahn u. a. „Fluid-Limiting Treatment Strategies among Sepsis Patients in the ICU: A Retrospective Causal Analysis“. In: *Critical Care* 24.1 (Feb. 2020), S. 62. ISSN: 1364-8535. DOI: 10.1186/s13054-020-2767-0.
- [Sha83] Ehud Y. Shapiro. „The Fifth Generation Project — a Trip Report“. In: *Commun. ACM* 26.9 (Sep. 1983), S. 637–641. ISSN: 0001-0782. DOI: 10.1145/358172.358179.
- [Sha92] Carla J Shatz. „The Developing Brain“. In: *SCIENTIFIC AMERICAN* (1992).
- [She+19] Jiayi Shen u. a. „Artificial Intelligence Versus Clinicians in Disease Diagnosis: Systematic Review“. In: *JMIR Med Inform* 7.3 (Aug. 2019), e10010. ISSN: 2291-9694. DOI: 10.2196/10010. PMID: 31420959.
- [She13] Henry M. Sheffer. „A Set of Five Independent Postulates for Boolean Algebras, with Application to Logical Constants“. In: *Transactions of the American Mathematical Society* 14 (1913), S. 481–488.
- [She26] Henry M. Sheffer. „Principia Mathematica. Whitehead, Alfred North , Russell, Bertrand“. In: *Isis; an international review devoted to the*

- history of science and its cultural influences* 8.1 (1926), S. 226–231. DOI: 10.1086/358383.
- [Sil10] R. Angus Silver. „Neuronal Arithmetic“. In: *Nature Reviews Neuroscience* 11.7 (Juli 2010), S. 474–489. ISSN: 1471-0048. DOI: 10.1038/nrn2864.
- [SKM10] K. Stromgaard, P. Krogsgaard-Larsen und U. Madsen, Hrsg. *Textbook of Drug Design and Discovery*. 4th ed. Detroit, MI: CRC Press, 2010. ISBN: 978-1-4200-6322-6.
- [Son22] Daniel Sonnet. „Neuronale Netze“. In: *Neuronale Netze Kompakt: Vom Perceptron Zum Deep Learning*. Wiesbaden: Springer Fachmedien Wiesbaden, 2022, S. 17–70. ISBN: 978-3-658-29081-8. DOI: 10.1007/978-3-658-29081-8_2.
- [SSH95] Peter Sander, Wolfried Stucky und Rudolf Herschel. „Automaten“. In: *Automaten Sprachen Berechenbarkeit: Grundkurs Angewandte Informatik IV*. Hrsg. von W. Stucky. Wiesbaden: Vieweg+Teubner Verlag, 1995, S. 26–97. ISBN: 978-3-322-84873-4. DOI: 10.1007/978-3-322-84873-4_2.
- [Ste22] Joachim Steinwendner. „Explainable AI Im Gesundheitswesen“. In: *Künstliche Intelligenz Im Gesundheitswesen: Entwicklungen, Beispiele Und Perspektiven*. Hrsg. von Mario A. Pfannstiel. Wiesbaden: Springer Fachmedien Wiesbaden, 2022, S. 755–767. ISBN: 978-3-658-33597-7. DOI: 10.1007/978-3-658-33597-7_36.
- [Str01] Bonnie B. Strickland, Hrsg. *The Gale Encyclopedia of Psychology*. 2nd ed. Detroit, MI: Gale Group, 2001. ISBN: 978-0-7876-4786-5.
- [Sze+15] Christian Szegedy u. a. *Rethinking the Inception Architecture for Computer Vision*. 2015. DOI: 10.48550/arXiv.1512.00567.
- [Szo88] Peter Szolovits. „Artificial Intelligence in Medical Diagnosis“. In: *Artificial Intelligence* 108.1 (1988). DOI: 10.7326/0003-4819-108-1-80.
- [TS81] Randy L. Teach und Edward H. Shortliffe. „An Analysis of Physician Attitudes Regarding Computer-Based Clinical Consultation Systems“. In: *Computers and Biomedical Research* 14.6 (1981), S. 542–558. ISSN: 0010-4809. DOI: 10.1016/0010-4809(81)90012-4.
- [Tur37] A. M. Turing. „On Computable Numbers, with an Application to the Entscheidungsproblem“. In: *Proceedings of the London Mathematical Society* s2-42.1 (1937), S. 230–265. DOI: 10.1112/plms/s2-42.1.230.
- [Uğü12] Harun Uğuz. „A Biomedical System Based on Artificial Neural Network and Principal Component Analysis for Diagnosis of the Heart Valve Diseases“. In: *Journal of Medical Systems* 36.1 (Feb. 2012), S. 61–72. ISSN: 1573-689X. DOI: 10.1007/s10916-010-9446-7.
- [Von11] H. Von Jepsen-Marwedel. „Einführung“. In: *Glastechnische Fabrikationsfehler: „Pathologische“ Ausnahmestände Des Werkstoffes Glas Und Ihre Behebung; Eine Brücke Zwischen Wissenschaft, Technologie Und Praxis*. Hrsg. von Hans Jepsen-Marwedel und Rolf Brückner. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, S. 3–15. ISBN: 978-3-642-16433-0. DOI: 10.1007/978-3-642-16433-0_1.

-
- [von93] J. von Neumann. „First Draft of a Report on the EDVAC“. In: *IEEE Annals of the History of Computing* 15.4 (1993), S. 27–75. DOI: 10.1109/85.238389.
- [Wie48] Norbert Wiener. *Cybernetics; or Control and Communication in the Animal and the Machine*. Oxford, England: John Wiley, 1948, S. 194.
- [WR27] Alfred North Whitehead und Bertrand Russell. *Principia mathematica, vol. 1 - 3*. 2. Aufl. Cambridge University Press, 1927. ISBN: 978-0521067911.
- [Wur09] Robert H. Wurtz. „Recounting the Impact of Hubel and Wiesel: Recounting the Impact of Hubel and Wiesel“. In: *The Journal of Physiology* 587.12 (Juni 2009), S. 2817–2823. ISSN: 00223751. DOI: 10.1113/jphysiol.2009.170209.
- [Zha+17] H. Zhao u. a. „Pyramid Scene Parsing Network“. In: *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Los Alamitos, CA, USA: IEEE Computer Society, Juli 2017, S. 6230–6239. DOI: 10.1109/CVPR.2017.660.

Abbildungsverzeichnis

2.1	Schematische Darstellung einer Nervenzelle. (Quelle: in Anlehnung an [SD07, 43, Tafel 2.1])	3
2.2	Ionenverteilung im Zytosol und der EZF. (Quelle: in Anlehnung an [BCP18, 72, Abb. 3.13])	5
2.3	Aktionspotenzial. (Quelle: in Anlehnung an [SD07, 47, Tafel 2.3])	7
3.1	Schematische Darstellung von MCP-Zellen (Quelle: in Anlehnung an [MP43, 105, Figure 1])	14
3.2	Plot einer Treppenfunktion, hier als Heaviside-Funktion mit $H(x) = \{x < 0 : 0, x \geq 0 : 1\}$. (Quelle: Eigene Darstellung)	15
3.3	Ein MCP-Neuron zur Modellierung von OR (Quelle: Eigene Darstellung)	18
3.4	Ein MCP-Netz zur Modellierung von NAND (Quelle: Eigene Darstellung)	19
3.5	Entwurf für ein MCP-Netz zur Modellierung von XOR (Quelle: Eigene Darstellung)	20
3.6	Rosenblatt-Perzeptron. (Quelle: in Anlehnung an [Roj93, 53, Abb. 3.2])	22
3.7	1. Winkelhalbierende im kartesischen Koordinatensystem (Quelle: Eigene Darstellung)	24
3.8	Halbräume im \mathbb{R}^2 (Quelle: Eigene Darstellung)	25
3.9	Punkte im \mathbb{R}^2 , die nicht durch eine Ursprungsgerade separierbar sind. Angedeutet eine mögliche Trenngerade, die durch $(0, 2)$ geht. (Quelle: Eigene Darstellung)	27
3.10	Perzeptron-Training für große Datenmengen. (Quelle: Eigene Darstellung)	28
3.11	XOR im \mathbb{R}^2 . (Quelle: Eigene Darstellung)	29
4.1	Exemplarische Darstellung der Architektur eines einfachen Feed-Forward-Netzes. (Quelle: Eigene Darstellung)	32
4.2	Plot einer Sigmoid-Funktion. (Quelle: Eigene Darstellung)	33
4.3	Skizzierung des Zusammenhangs Fehlerfunktion und lokale Minima (Quelle: in Anlehnung an [Son22, 52, Abb. 2.15])	34

4.4	Beispiel einer Faltungsoperation. (Quelle: in Anlehnung an [GBC18, 372, Abbildung 9.1])	38
4.5	Plot der ReLU. (Quelle: Eigene Darstellung)	38
A.1	Phasen des Aktionspotenzials. (Quelle: in Anlehnung an [BCP18, 88, Abb. 4.1b])	51
C.1	MCP-Netz zur Modellierung der paradoxen Hitzeempfindung. (Quelle: in Anlehnung an [Fau94, 31, Figure 1.18])	68
C.2	Zwei Kälteimpulse als Eingabe für das MCP-Netz. (Quelle: in Anlehnung an [Fau94, 35, Figure 1.23]).	69
C.3	Ein einzelner Wärmeimpuls aktiviert das Wärmesignal. (Quelle: in Anlehnung an [Fau94, 36, Figure 1.24]).	70
C.4	Entstehung der “paradoxen Hitzeempfindung“ durch einen einzelnen Kälteimpuls. (Quelle: in Anlehnung an [Fau94, 34, Figure 1.22]).	71
D.1	Geometrische Interpretation der Perzeptron-Funktion im \mathbb{R}^2 (Quelle: Eigene Darstellung)	72
D.2	Geometrische Darstellung der Konvergenz eines Perzeptrons für AND (Quelle: Eigene Darstellung)	79