

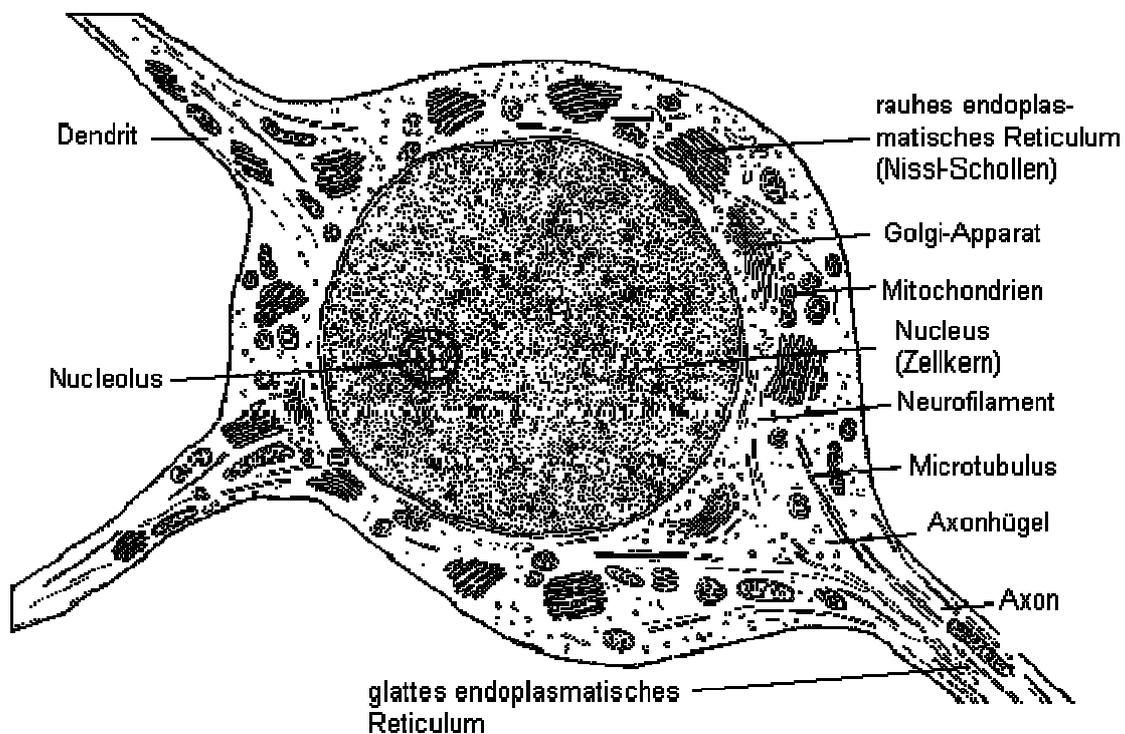
Neuronale Netze

Während Probleme, die durch einen Algorithmus in kurzer Zeit exakt zu lösen sind, von einem Computer deutlich schneller gelöst werden können als von einem Menschen, benötigt das menschliche Gehirn für Aufgaben wie das Erkennen eines Gesichts wesentlich weniger Zeit. Außerdem erreicht der Computer nicht die hohe Erkennungsleistung. Ein weiterer Vorteil des menschlichen Gehirns ist, daß auch dann noch korrekte Ergebnisse geliefert werden, wenn es zu einem Ausfall einiger für die Problemlösung notwendiger Nervenzellen kommt. Selbst wenn die "Eingaben" ungenau sind, also beispielsweise ein Text durch Verschmutzung unleserlich geworden ist, kann das Gehirn den Text noch erkennen. Ein Computer liefert in diesen Fällen fehlerhafte bzw. unbrauchbare Ergebnisse. Die Idee ist daher, die Arbeitsweise des Gehirns auf Maschinen zu übertragen.

Künstliche neuronale Netze bestehen - wie das Gehirn von Säugetieren - aus einer großen Anzahl kleiner Elemente, den Neuronen. Information wird verarbeitet, indem sich die Neuronen mit Hilfe von gerichteten Verbindungen untereinander aktivieren. Dies geschieht im Prinzip analog zu den Vorgängen im Gehirn.

Neuronale Netze zeichnen sich durch ihre Lernfähigkeit aus. Sie können eine Aufgabe anhand von Trainingsbeispielen erlernen, ohne dazu explizit programmiert werden zu müssen. Weitere Vorteile sind die hohe Parallelität bei der Informationsverarbeitung, die hohe Fehlertoleranz und die verteilte Wissensrepräsentation, wodurch ein zerstörtes Neuron nur einen relativ kleinen Wissensausfall bedeutet.

Aufbau einer Nervenzelle

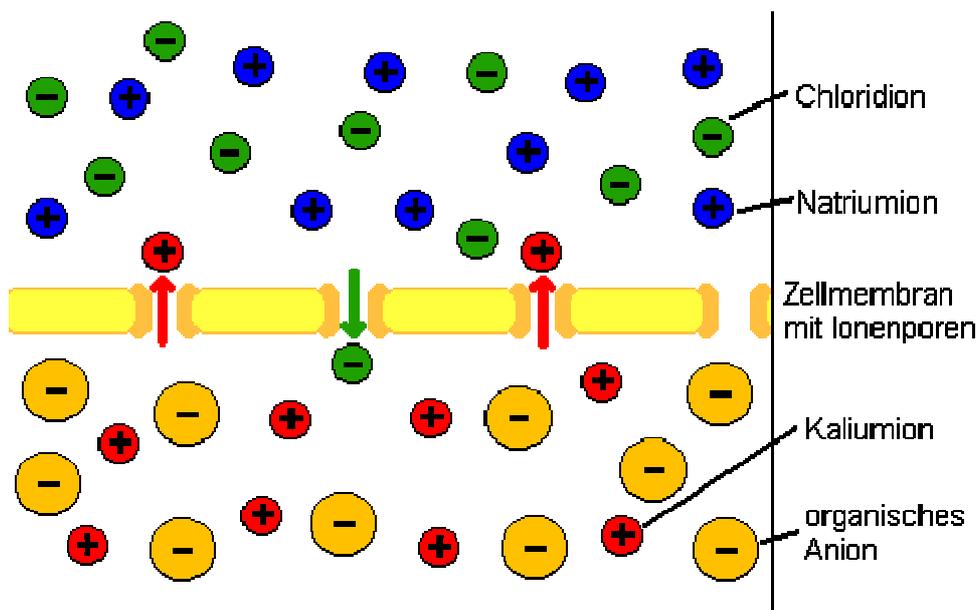


Nervenzellen oder Neuronen sind die Grundbausteine des Nervensystems. Sie bestehen aus einem bis zu 0,25mm großen Zellkörper, aus dem kurze verästelte Fortsätze, die Dendriten, hervorgehen. Ebenso wie andere Zellen besitzen sie Organellen, welche die Zellen mit Energie und anderen notwendigen Stoffen versorgen. Dazu gehören auch das **endoplasmatische Reticulum** und der **Golgi-Apparat**. Diese Organellen produzieren die für die Weiterleitung von Nervensignalen notwendigen synaptischen Vesikel, welche die Neurotransmitter enthalten. Im Gegensatz zu anderen Zellen bilden Nervenzellen Ausläufer aus, die an den Enden Verdickungen, die sogenannten **Synapsen** haben können. Diese sorgen für die **Weiterleitung von Reizen** an andere Nerven- oder Muskelzellen. Sie enthalten die synaptischen Vesikel mit den Neurotransmittern, die der Übertragung von Reizen von einer Zelle auf die andere dienen.

Einige Ausläufer der Neuronen, die sogenannten **Dendriten**, nehmen Eingangssignale auf. Sie sind stark verästelt. Das **Axon**, die normalerweise viel längere Nervenfasern, dient dagegen der Weiterleitung von Nervenreizen an andere Zellen. Axone können sich an ihrem Ende verästelten und besitzen dort die Synapsen für die Reizweiterleitung.

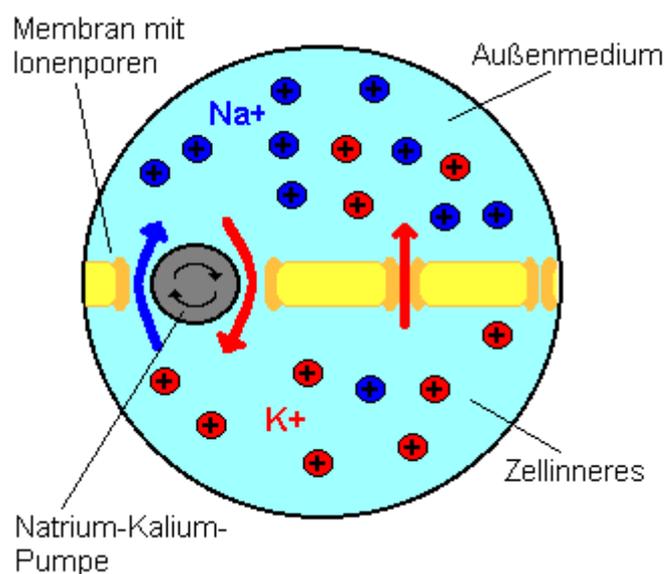
Arbeitsweise von Nervenzellen

Die ca. 5nm dicke **Zellmembran** einer Nervenfaser besteht aus einer Doppelschicht von **Lipiden**, deren fettliebende (lipophile) Seite nach innen und deren wasserliebende Seite nach außen zeigt. In diese **Lipiddoppelschicht** sind Proteine (Eiweiße) eingebaut, die verschiedene Funktionen erfüllen. Bei einigen dieser Proteine handelt es sich um **Ionenkanäle**, welche die Zellmembran für bestimmte Ionen durchlässig machen.



Die **Ionenkonzentrationen** innerhalb und außerhalb der Zelle sind unterschiedlich. Die Konzentration der positiv geladenen Kaliumionen (K^+) ist im Zellinneren höher als im Außenmedium, positiv geladene Natriumionen (Na^+) kommen dagegen vermehrt außerhalb der Zelle vor. Dieses Konzentrationsgefälle wird durch die sogenannte **Ionenpumpe** aufrechterhalten, die in einem Pumpzyklus 3 Natriumionen nach außen und 2 Kaliumionen in die Zelle hinein bringt. Die dazu benötigte Energie wird durch die Spaltung von ATP (Adenosintriphosphat) gewonnen.

Während nun Kalium- und negativ geladene **Chloridionen** (Cl^-) durch die Ionenkanäle diffundieren können, ist die Zellmembran für **Natriumionen** praktisch undurchlässig. Wegen ihrer hohen Konzentration im Zellinneren strömen **Kaliumionen** laufend durch die Membranporen ins Außenmedium. Chloridionen, deren Konzentration im Außenmedium höher ist, diffundieren in geringerem Umfang ins Innere der Zelle. Dadurch entsteht außerhalb der Zelle ein Überschuß an positiv geladenen, innen dagegen ein Überschuß an negativ geladenen Ionen. Durch diese Potentialdifferenz wird der Ausstrom der Kaliumionen gebremst, so daß sich ein Gleichgewichtszustand zwischen der nach außen gerichteten Diffusionstendenz und der nach innen gerichteten elektrischen Anziehung für Kaliumionen einstellt. Umgekehrt gilt dies für die Chloridionen. Es resultiert ein Gleichgewichtszustand, bei dem das Zellinnere gegenüber dem Zelläußeren negativ geladen ist. Dieser Gleichgewichtszustand wird als **Ruhepotential** bezeichnet.



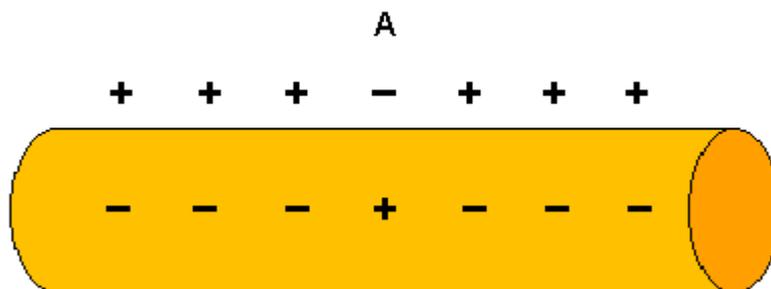
Durch geringen Ausstrom von Kaliumionen entsteht das **Ruhepotential**. Das Zellinnere ist gegenüber dem Außenmedium negativ geladen.

In der Membran eines Axons befinden sich **spannungsgesteuerte Porenproteine**, deren Durchlässigkeit vom Membranpotential abhängig ist. Beim Ruhepotential sind die Natriumkanäle geschlossen, nur ein Teil der Kaliumporen ist geöffnet. Wird das Axon durch einen elektrischen Reiz etwas depolarisiert, öffnen sich einige Natriumporen. Erst wenn die Depolarisation einen Schwellenwert erreicht, öffnen sich alle Natriumkanäle, während die Anzahl der durchlässigen Kaliumporen zunächst unverändert bleibt. Deshalb strömen zu Beginn des **Aktionspotentials** mehr Natriumionen nach innen als Kaliumionen nach außen, so daß im Inneren des Axons ein Überschuß an positiver Ladung entsteht. Die Membran ist also gegenüber dem Ruhezustand umgekehrt geladen. Nach 1-2ms schließen sich die Natriumporen wieder, während sich alle noch geschlossenen Kaliumkanäle öffnen. Durch einen erhöhten Ausstrom an Kaliumionen kehrt das Membranpotential rasch wieder zum Ruhewert zurück.

Fortpflanzung des Nervensignals

Die **Weiterleitung des Aktionspotentials** erfolgt so: Entsteht durch Reizung einer bestimmten Stelle ein Aktionspotential, grenzen an dieser Stelle positive und negative Ladungen ohne trennende Membran aneinander. Es entstehen Ausgleichsströme, da sich Ionen unterschiedlicher Ladung gegenseitig anziehen. Das Membranpotential der benachbarten Stellen wird auf diese Weise erniedrigt. Ist der Schwellenwert erreicht, entsteht auch hier ein Aktionspotential. Diese neu entstandenen Aktionspotentiale erniedrigen ihrerseits wieder das Aktionspotential benachbarter Stellen. Dies geschieht auch an der ursprünglichen Reizstelle, die jedoch noch unerregbar ist, so daß hier kein Aktionspotential entsteht.

Nervenfaser, bei A gereizt

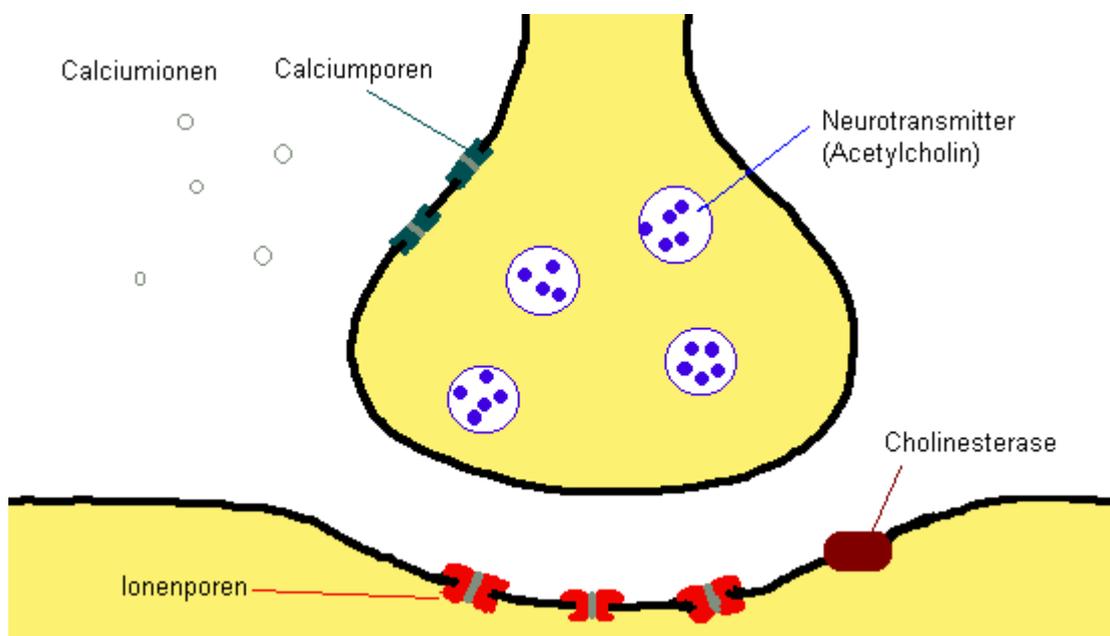


Wichtig für das Verständnis der Weiterleitung des Nervensignals ist, daß das Aktionspotential **immer** die **gleiche Amplitude** und die **gleiche Form** besitzt. Das bedeutet, daß ein einzelnes Signal außer durch den Ort, an dem es gesendet wird, keine Informationen tragen kann. Die **Information** wird also **über die Frequenz** der Aktionspotentiale sowie **die Dauer der Entstehung** von Aktionspotentialen codiert.

Interessant ist auch, daß sich die Eigenschaften einer Synapse kurz- oder längerfristig ändern können. Bei wiederholter Reizung kann eine verminderte Ausschüttung des Neurotransmitters erfolgen. Dies bezeichnet man als **Habituation** (Gewöhnung). Andererseits kann die Ausschüttung der Transmitter auch erhöht werden, also eine **Sensibilisierung** erfolgen. Während Habituation und Sensibilisierung nur verhältnismäßig kurze Zeit anhalten, gibt es auch längerfristige Veränderungen der Aktivität von Synapsen, die durch strukturelle Veränderungen der Synapsenregion entstehen.

Signalübertragung an den Synapsen

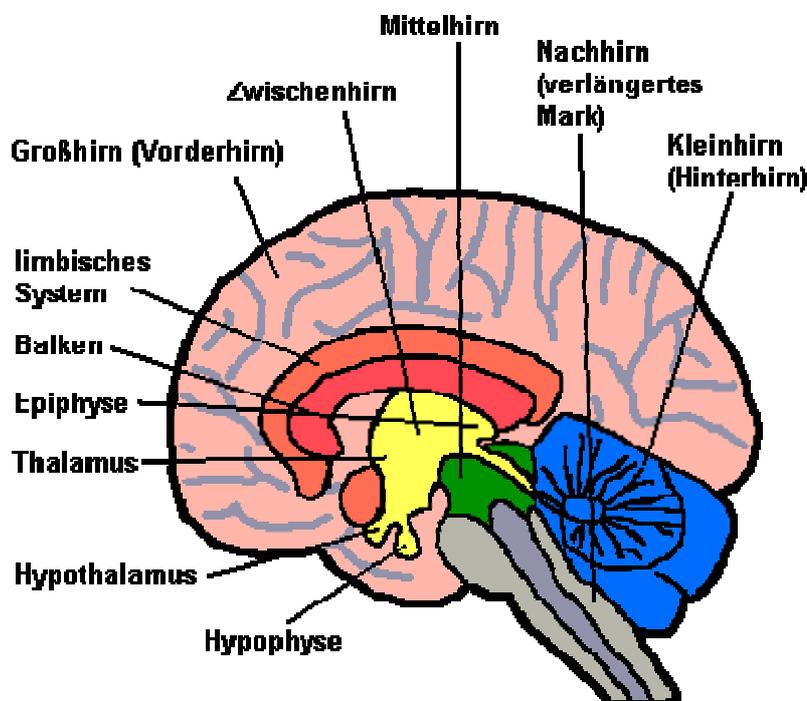
Zu betrachten ist noch die **Signalübertragung an den Synapsen**. Zwischen dem Endknopf eines Axons und der Membran der anschließenden Muskel- oder Nervenzelle befindet sich ein schmaler, mit Flüssigkeit gefüllter Spalt, der 20nm breite, sogenannte synaptische Spalt. Man unterscheidet an einer Synapse den sogenannten präsynaptischen, also vor dem Spalt liegenden Teil und den postsynaptischen Teil, der sich hinter dem Spalt befindet. In der Membran des Endknopfes befinden sich Calciumkanäle, die in Ruhe geschlossen sind. Das Zellinnere ist dann relativ arm an Calciumionen (Ca^{2+}).



Erreicht ein Aktionspotential den Endknopf, öffnen sich die Calciumporen kurzzeitig, so daß die Ca^{2+} -Konzentration im Inneren der Zelle ansteigt. Dadurch wird bewirkt, daß ein Teil der synaptischen Bläschen mit der Zellmembran verschmilzt und der enthaltene Neurotransmitter Acetylcholin in den synaptischen Spalt gelangt. Die Muskelfaser hat wie die Nervenfasern ein Ruhepotential. Durch das Andocken von Acetylcholin an Rezeptoren der postsynaptischen Membran öffnen sich Ionenkanäle. Dadurch können Natriumionen ein- und Kaliumionen - in geringerem Umfang - ausströmen. Die Differenz zwischen dem Ruhepotential und dem erniedrigten Membranpotential bezeichnet man als Endplattenpotential. Die Acetylcholin-Moleküle bewegen sich im synaptischen Spalt wie Pingpongbälle hin und her und können so mehrere Ionenkanäle hintereinander öffnen. Erreicht das Endplattenpotential einen Schwellenwert, löst es in der postsynaptischen Muskel- bzw. Nervenzelle ein Aktionspotential aus. Um eine Dauererregung zu verhindern, werden die Acetylcholin-Moleküle von einem Enzym, der Cholinesterase, gespalten und so unwirksam gemacht.

Das Nervensystem

Eine leistungsfähige Informationsverarbeitung kann erst durch das Zusammenwirken vieler Nervenzellen in einem Nervensystem erfolgen. Das Nervensystem verarbeitet die von den Sinnesorganen ankommenden Informationen und kontrolliert die Tätigkeit der Muskeln und Drüsen des Körpers. Bei allen Wirbeltieren liegt das Zentralnervensystem auf der Rückenseite des Körpers. Es besteht aus Gehirn und Rückenmark. Allein das Gehirn des Menschen enthält etwa 13 Millionen Nervenzellen. Das periphere Nervensystem stellt die Verbindung zu den inneren Organen her (vegetatives Nervensystem) und versorgt die Skelettmuskulatur (somatisches Nervensystem).



Das Vorder- oder Großhirn besteht aus zwei Hälften. Die motorischen Regionen steuern die willkürlichen Bewegungen der Skelettmuskeln, indem sie ihre Befehle an die zuständigen Regionen des Rückenmarks übermitteln. In die sensorischen Regionen gelangen die Informationen aus den Sinnesorganen. Die Assoziationsregionen verknüpfen Meldungen der Sinnesorgane untereinander und mit Informationen aus anderen Gehirnteilen. Sie sind verantwortlich für alle höheren Leistungen des Gehirns. Von besonderer Bedeutung ist das limbische System, bei dem es sich um eine Art vegetatives Gehirn handelt. Es ist außerdem wesentlich an der Entstehung der verhaltenssteuernden Triebe beteiligt. Auch für Gedächtnis und Lernfähigkeit sowie sprachliche Leistungen und Bewußtsein ist das Großhirn zuständig.

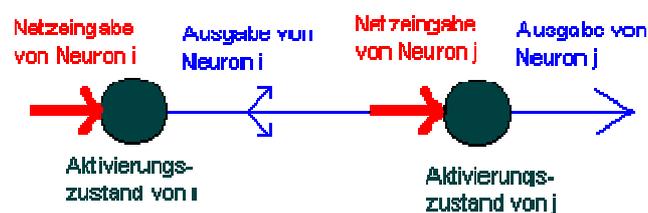
Die wichtigsten Teile des Zwischenhirns sind der Thalamus und der Hypothalamus. Während der erste Teil bei Säugetieren die Hauptumschaltstelle zwischen den Sinnesorganen und dem Großhirn ist, steuert der Hypothalamus das vegetative Nervensystem. Er sorgt auch für die Regelung der Körpertemperatur.

Das Mittelhirn ist bei niederen Wirbeltieren die Hauptumschaltstelle zwischen den Sinnesorganen und der Muskulatur, bei Säugetieren hat es in dieser Hinsicht nur eine untergeordnete Bedeutung. Es ist auch für die Steuerung der Höhe der Bewußtseinslage verantwortlich (Schlaf, Bewußtlosigkeit). Das Hinter- oder Kleinhirn spielt eine große Rolle für das Gleichgewicht. Es ist jederzeit über die Stellung des Körpers im Raum und die Lage der einzelnen Gliedmaßen zueinander informiert. Eine Ausschaltung des Kleinhirns macht es unmöglich, schnell aufeinanderfolgende Bewegungen auszuführen. Die Betroffenen gehen taumelnd und wirken ungenau, die Koordination der Bewegungen ist gestört. Das Nachhirn ist die Übergangsstelle zwischen Rückenmark und Gehirn. Es ist nicht nur Schaltstelle aller vom Gehirn zum Rückenmark und umgekehrt laufender Bahnen, sondern auch Sitz vieler Zentren für lebenswichtige Reflexe. Außerdem liegt hier das Zentrum für die Regelung des Blutdrucks.

Modellierung von Neuronen

Ein künstliches neuronales Netz besteht aus stark idealisierten Neuronen. Wie ihr biologisches Vorbild bestehen sie aus drei Komponenten: dem **Zellkörper**, den **Dendriten** und einem **Axon**.

Die Dendriten summieren die Eingabe des Netzes in die Zelle auf, das Axon leitet die Ausgabe der Zelle an die Dendriten nachfolgender Synapsen weiter. Die Stärke der Synapsen wird durch einen numerischen Wert, das **Verbindungsgewicht**, dargestellt. Daher lässt sich die Verbindung zwischen Neuronen als direkte gewichtete Verbindung zwischen den beiden Zellen **i** und **j** darstellen.



Komponente		Eigenschaften und Aufgaben
Zellen (Neuronen)		
	Aktivierungs- zustand Z	gibt den Aktivierungsgrad der Zelle an
	Aktivierungs- funktion f_{act}	gibt an, wie sich ein neuer Aktivierungszustand $Z_{neu,j}$ des Neurons j aus der alten Aktivierung $Z_{alt,j}$ und der Netzeingabe $net_i(t)$ sowie dem Schwellenwert des Neurons j ergibt.
	Ausgabe- funktion f_{out}	bestimmt aus der Aktivierung die Ausgabe des Neurons
Verbindungs- netzwerk der Zellen		Ein neuronales Netz kann als gerichteter, gewichteter Graph angesehen werden. Die Kanten stellen die Verbindungen zwischen den Neuronen dar. $w_{i,j}$ ist das Gewicht (weight) der Verbindung von Neuron i nach Neuron j , die Matrix W aller Verbindungen heißt Gewichtsmatrix.
Propagie- rungs- funktion		gibt an, wie sich die Netzeingabe eines Neurons aus den Ausgaben der anderen Neuronen und den Verbindungsgewichten berechnet. Dabei handelt es sich um die gewichtete Summe der Ausgaben der Vorgängerzellen.
Lernregel		Algorithmus, nach dem das Netz lernt, für eine vorgegebene Eingabe eine gewünschte Ausgabe zu produzieren. Durch die wiederholte Eingabe von Trainingsmustern wird die Stärke der Verbindungen zwischen den Neuronen modifiziert. Dabei wird versucht, den Fehler zwischen erwarteter und tatsächlicher Ausgabe des Netzes zu minimieren. Lernverfahren sind die interessanteste Komponente der neuronalen Netze.

Definition:

Ein künstliches Neuron ist ein Tupel (x, w, f_a, f_o, o) bestehend aus einem Eingabevektor $x = (x_1, \dots, x_n)$, einem Gewichtsvektor $w = (w_1, \dots, w_n)$, einer Aktivierungsfunktion f_a mit $f_a: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ und einer Ausgabefunktion f_o , für die $f_o: \mathbb{R} \rightarrow \mathbb{R}$ gilt.

Dabei wird durch $f_o(f_a(x, w))=o$ der Ausgabewert des Neurons erzeugt, der an die nachfolgenden Neuronen über die Axonkollaterale weitergeleitet wird.

Faßt man Aktivierungs- und Ausgabefunktion zu einer **Transfer-Funktion** zusammen, verwendet man häufig die gewichtete Summe:

$$f_a(\vec{x}, \vec{w}) = \sum_{k=i}^n x_k w_k$$

Außerdem kann man dem einzelnen Neuron, analog zur biologischen Nervenzelle, einen inneren Aktivierungszustand zuordnen. Die Definition für ein Neuron kann dann auch so aussehen:

Definition:

Ein künstliches Neuron ist ein Tupel (x, w, f_a, Z, f_o, o) bestehend aus:

1. Eingabevektor x
2. Gewichtsvektor w
3. Aktivierungsfunktion f_a mit $f_a: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$
4. Zustand Z
5. Ausgabefunktion $f_o: \mathbb{R} \rightarrow \mathbb{R}$

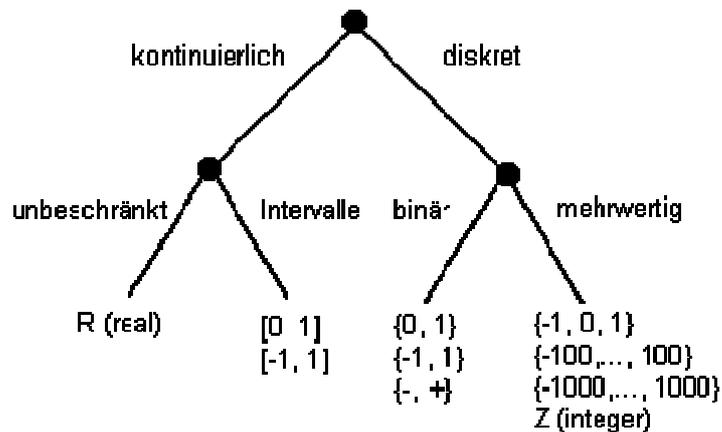
Hierbei ist die Ausgabefunktion f_o eine Funktion in Abhängigkeit von Z .

Der **Aktivierungszustand** Z hängt vom alten Zustand und der Veränderung der Aktivierungsfunktion ab und kann beispielsweise durch

$$Z^{neu} = Z^{alt} \mid f_a(\vec{x}, \vec{w})$$

ausgedrückt werden.

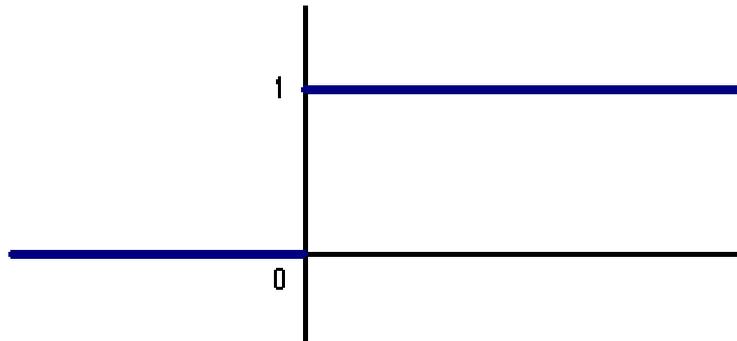
In der technischen Realisierung wird er unterschiedlich dargestellt. Man unterscheidet **(quasi-)kontinuierliche und diskrete Wertebereiche**, die sich wiederum in unterschiedliche Modelle aufteilen lassen:



Bei kontinuierlichen Wertebereichen beschränken die meisten Modelle den [Aktivierungszustand](#) auf ein Intervall. Das liegt daran, daß es sich meistens um eine nichtlineare, häufig sigmoide, [Aktivierungsfunktion](#) und die Identität als Ausgabefunktion handelt. Dadurch wird die Ausgabe identisch mit der Aktivierung, und der Wertebereich der Aktivierungsfunktion gibt den Wertebereich des Aktivierungszustandes an. Das ursprüngliche Hopfield-Modell und andere Modelle verwenden diskrete Aktivierungszustände. Diese werden dann auch in der Implementierung als binäre Werte gespeichert und verarbeitet. Damit eine biologische Nervenzelle "feuert", also ein Aktionspotential ausgelöst wird, muß ein bestimmter [Schwellenwert](#) S überschritten werden.

Auch für künstliche Neuronen gibt es eine **Schwellenwertfunktion**:

$$f_o\left(\sum_{k=1}^n x_k w_k\right) = \begin{cases} 1: & \text{falls } \sum_{k=1}^n x_k w_k \geq S \\ 0: & \text{sonst} \end{cases}$$

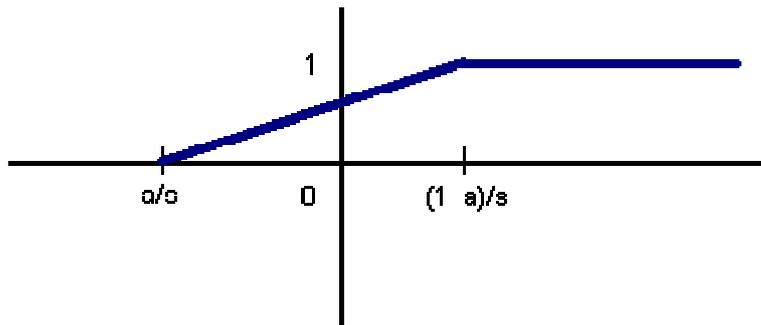


Da diese Art der **Ausgabefunktion** nicht die Intensität der aufeinanderfolgenden Aktionspotentiale eines biologischen Neurons simuliert, werden lineare Ausgabefunktionen verwendet. Der zeitliche Abstand, in dem die Aktionspotentiale durch die biologische Nervenzelle

weitergereicht werden, ist nach unten beschränkt. Daher sollte auch im künstlichen Neuronenmodell eine beschränkte Ausgabefunktion Verwendung finden.

Diese Ausgabefunktionen lassen sich durch semilineare Funktionen der Art

$$f_o\left(\sum_{k=1}^n x_k w_k\right) = \begin{cases} 1 & : \text{falls } \sum_{k=1}^n x_k w_k \geq \frac{1+a}{s} \\ s\left(\sum_{k=1}^n x_k w_k\right) - a & : \text{falls } \frac{a}{s} \leq \sum_{k=1}^n x_k w_k < \frac{1+a}{s} \\ 0 & : \text{sonst} \end{cases}$$



beschreiben.

Sinnvoller ist es aber, die Aktivierung beziehungsweise die Ausgabe durch glattere, also differenzierbare Funktionen zu beschreiben. Ein Beispiel für differenzierbare und beschränkte Funktionen sind die s-förmigen oder sigmoiden Funktionen:

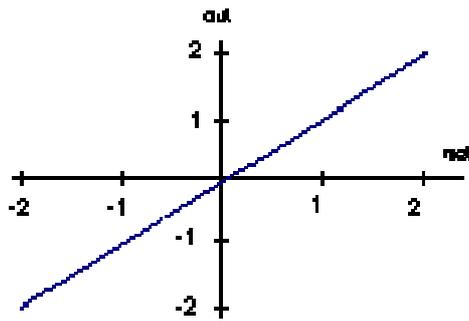
Definition:

Eine Funktion $s_c: \mathbb{R} \rightarrow [0, 1]$ heißt **sigmoide oder s-förmige Funktion**, wenn sie monoton wachsend und differenzierbar ist und wenn

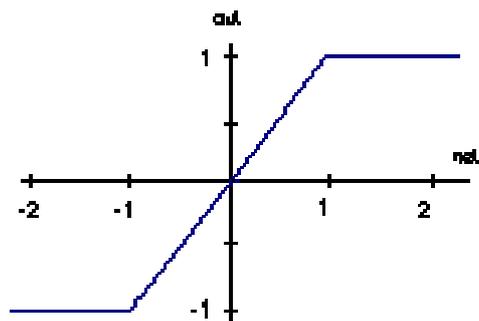
$$\lim_{\lambda \rightarrow \infty} s_c(\lambda) = k_1 \quad \text{und} \quad \lim_{\lambda \rightarrow -\infty} s_c(\lambda) = k_2 \quad \text{mit} \quad k_2 < k_1$$

gelten.

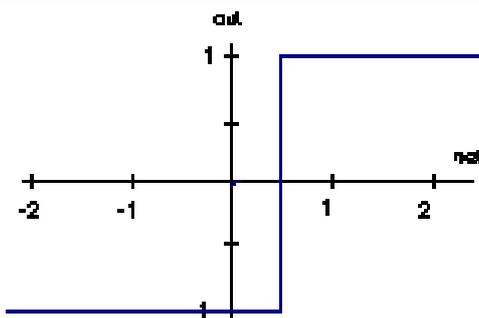
Die folgende Abbildung enthält einige gebräuchliche Aktivierungs- bzw. Ausgabefunktionen, wobei die beiden letzteren am häufigsten verwendet werden.



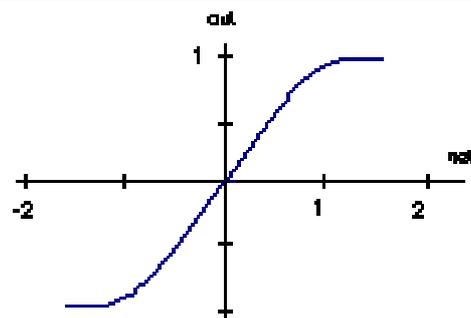
Identität ($o_j = net_j$)



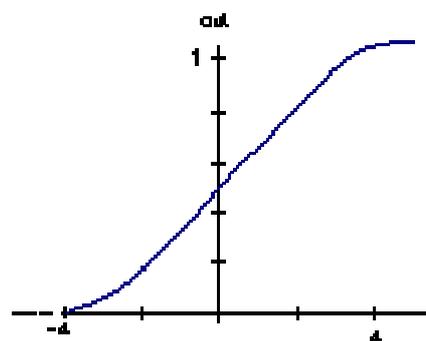
linear bis Sättigung



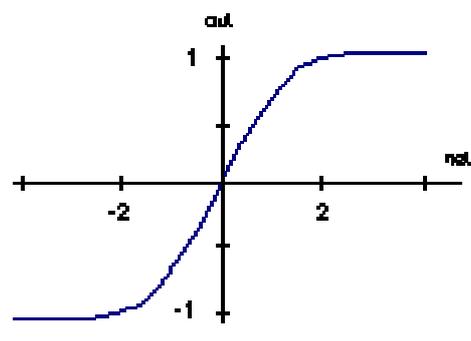
binäre Schwellenwertfkt.



$\sin(x)$ bis Sättigung



$(1 / (1 + \exp(-x)))$



$\tanh(x)$