

**Max-Planck-Institut
für Mathematik
in den Naturwissenschaften
Leipzig**

**Kognitive Modellierung mit künstlichen
neuronalen Netzen**

by

Thomas Wennekers and Günther Palm

Preprint no.: 27

2000



Kognitive Modellierung mit künstlichen neuronalen Netzen*

Thomas Wennekers¹ und Günther Palm²

¹ Max-Planck-Institut für Mathematik in den Naturwissenschaften,
Inselstraße 22-26, 04103 Leipzig
Thomas.Wennekers@mis.mpg.de

² Abteilung Neuroinformatik, Universität Ulm, 89069 Ulm
Palm@neuro.informatik.uni-ulm.de

1 Einleitung

Auf den ersten Blick steckt in dem Thema dieses Aufsatzes ein ungeheurer Anspruch, nämlich die Modellierung kognitiver Leistungen auf der Basis der Neurone, also der vielen einzelnen Nervenzellen im Gehirn. Das wäre ein Traumziel der Hirnforschung, eine vollständige Reduktion kognitiver Leistungen auf die physikalischen Prozesse in den Neuronen und Synapsen.

In der Tat sind wir heute noch weit davon entfernt. Auf der einen Seite werden die experimentellen Forscher in der Neurobiologie nicht müde, uns immer wieder mit neuen Details zu überraschen, die zeigen, daß die physikalischen Prozesse in den Neuronen und Synapsen noch lange nicht vollständig verstanden sind. Auf der anderen Seite ist es bis heute schlicht unmöglich, selbst stark vereinfachte Modelle von Milliarden von Neuronen in vernünftiger Zeit auf einem Computer zu simulieren. Selbst das Festlegen der Parameter und der Startwerte wäre schon nicht ganz unproblematisch. Wie nähert man sich also der kognitiven Modellierung mit künstlichen neuronalen Netzen? Es gibt hier prinzipiell drei Möglichkeiten.

1. Man modelliert einfache Nervensysteme (etwa den Plattwurm) oder Teilsysteme

*To appear in: T.Goschke, M.Eimer, eds., Enzyklopädie der Psychologie, Band C/II/8 - Kognitive Neurowissenschaft. Hogrefe Verlag, Göttingen, 2000.

(etwa das Rückenmark). Dies widerspricht allerdings der landläufigen Vorstellung von wirklich kognitiven Leistungen.

2. Man modelliert eine spezielle einzelne kognitive Leistung, die nur einen kleinen Teil des Gehirns benötigt. Hier ist das Problem, überhaupt eine solche Leistung zu finden.
3. Man modelliert nicht wirklich einzelne Neurone, sondern die “Einheiten” oder “Units” des Modells stehen bereits für größere “Gruppen”, “Pools”, “Kolumnen” oder “Assemblies” dieser größeren Entitäten.

Wir wollen uns im folgenden besonders auf den dritten Weg konzentrieren. Dies paßt auch gut zum heute gängigen Gebrauch der “künstlichen neuronalen Netze” oder auch der “konnektionistischen Modelle” als Modellierungsinstrument. In der Technik ist die detailgetreue Ähnlichkeit zwischen künstlichen und echten biologischen Neuronen sowieso unwichtig (Schöneburg, 1993), Hauptsache die Netze erfüllen ihre Funktion in der Mustererkennung (Bishop, 1995; Ripley, 1996; Schürmann, 1996), bei der Vorhersage von Börsenkursen (Bol et al., 1995), oder bei der Steuerung komplexer Anlagen (Albus, 1975; 1991; 1992; Tolle und Ersü, 1992; Wolpert, 1997). In einem Zweig der Kognitionsforschung werden künstliche neuronale Netze (dort zumeist unter Berufung auf Rumelhart und McClelland (1986) konnektionistische Modelle genannt) tatsächlich zur Modellierung kognitiver Funktionen eingesetzt, aber die einzelnen “Units” in dem Netz stellen üblicherweise nicht einzelne Neurone dar und müssen diesen deswegen auch nicht detailgetreu nachgebildet sein (vgl. auch Feldman und Ballard, 1982; McClelland et al., 1986). Einige Beispiele für den Einsatz künstlicher neuronaler Netze in diesem Bereich findet man etwa bei der Modellierung der visuellen Aufmerksamkeit (Koch and Ullmann, 1985; Itti et al., 1998; Suder et al., 2000), der kognitiven Karten (Recce and Harris, 1996; Harris and Recce, 1997; Mallot, 1997), der Produktion von Sprache (Cohen und Grossberg, 1986; Sejnowski und Rosenberg, 1987) oder dem Sprachverstehen (Lippmann, 1989; Pulvermüller und Preißl, 1991; Pulvermüller, 1992; Miikkulainen, 1993; McQueen et al., 1995; Wermter, 1995).

Welche Eigenschaften sind es nun, die für solche Modelle die Bezeichnung “neuro-nal” noch rechtfertigen? Oder: Was sind die minimalen strukturellen Eigenschaften von wirklichen Neuronen, die in solche Modelle Eingang finden bzw. Eingang finden sollten?

Die Grundidee ist wohl, daß die Komplexität der Verarbeitung nicht aus den Eigenschaften der Einzelelemente entsteht (die also ruhig möglichst simpel sein sollten), sondern aus deren Verschaltung in einem Netzwerk. Wichtig sind also eher Verschaltungsprinzipien, nämlich:

1. Hohe Konvergenz und Divergenz: Jedes Neuron enthält von vielen anderen Eingänge und gibt seinen Ausgang an viele andere weiter.
2. Rückkopplung: Es gibt in dem Verschaltungsgraphen des neuronalen Netzes sehr viele Zyklen.

3. Lernen durch veränderliche Verbindungen: Die Veränderung der “Stärke” einer Synapse wird auch als synaptische Plastizität bezeichnet. Diese wurde experimentell im Hinblick auf die zugrundeliegenden biochemischen Mechanismen und in der Theorie im Hinblick auf funktionelle Konsequenzen untersucht. Die gängigen Ideen sind allesamt Varianten der sogenannten Hebb Synapsen (Hebb, 1949; Bienenstock et al., 1982; Artola und Singer, 1993; Bliss und Collingridge, 1993; Abbott und Blum, 1996; Palm, 1982ab; u.a.).

Diese Prinzipien sind in den meisten konnektionistischen Modellen kognitiver Leistungen nicht erfüllt. Wenn man allerdings die Forderungen (1) und (2) hinsichtlich des quantitativen Aspektes entschärft und schon mit gelegentlicher Konvergenz, Divergenz und Rückkopplung zufrieden ist, dann gibt es bereits viele akzeptable Modelle (Elman, 1990; Pulvermüller, 1998). Wenn man auch noch auf die Rückkopplung verzichtet und die neuronal etwas schwer erklärliche Lernregel des Backpropagation-Verfahrens (Werbos, 1988; Rumelhart und McClelland, 1996) akzeptiert, kommt eine große Zahl weiterer Modelle hinzu (zu viele, um hiervon einzelne noch einmal durch Zitate hervorzuheben). Eine Übersicht vermitteln Lehrbücher wie z.B. Rojas (1993), Haykin (1994), Brause (1995) und Bishop (1995).

Betrachten wir also zunächst Modelle kognitiver Leistungen, welche die Punkte 1 bis 3 zwar nicht quantitativ, aber qualitativ einigermaßen erfüllen. Eine interessante Frage an diese Modelle ist die Interpretation der “Units”. Was wurde hier als die “neuronale Einheit” gewählt?

In dieser Frage gibt es zwei bis drei naheliegende Ideen: Offensichtlich kann man solche Neurone zusammenzufassen, die dasselbe tun oder die sich zumindest ähnlich verhalten. Das bedeutet: die zusammengefaßten Neurone sollten ungefähr das gleiche Input-Output-Verhalten haben und sie sollten in etwa die gleichen Eingangsreize bekommen. Solche Gruppen werden oft als “Assemblies”, “Ensembles” oder einfach “Pools” bezeichnet. Hinzu kommen zwei weitere naheliegende Forderungen: (1) Die zusammengefaßten Neurone sollten untereinander massiv exzitatorisch verbunden sein (eine Forderung die bei Assemblies meistens gestellt wird). (2) Sie sollten räumlich benachbart sein; diese Gruppen werden oft als “Kolumnen” bezeichnet.

Hinter der letzten Forderung steht die Vorstellung, daß räumlich benachbarte Neuronen sich ähnlich verhalten und ähnliche Eingänge bekommen. Gerade in der Großhirnrinde spricht man dann von Kolumnen oder Säulen aus tangential benachbarten Neuronen, die sich in radialer Richtung durch sämtliche Cortexschichten erstrecken. Dieser Vorstellung (der Einheitlichkeit der Neurone einer Kolumne) wird von Seiten der Experimentatoren im Detail immer wieder heftig widersprochen, weil man leicht unmittelbar benachbarte Neurone mit ganz verschiedenen Antworteigenschaften finden kann. Andererseits ist diese Vorstellung eng verbunden mit der bei Experimentatoren eher beliebten Vorstellung von “Karten”, in denen bestimmte Antworteigenschaften in kontinuierlicher Weise auf dem Cortex abgebildet sind, also so, daß sich benachbarte Neurone hinsichtlich dieser Eigenschaften ähnlich verhalten. Alles in allem läßt sich der Kolumnen- oder Kartengedanke

wohl am ehesten in den primären oder sekundären sensorischen oder motorischen Arealen aufrechterhalten. Insgesamt wäre es allerdings eine Verschwendung von Ressourcen, wenn durchgängig überall im Hirn beispielsweise die vielen Millionen Neuronen innerhalb eines Kubikmillimeters fast dasselbe tun würden. Solche Effizienzüberlegungen sind jedoch in dem von uns entwickelten Assembly-Konzept, das auf Donald Hebb (1949) zurückgeht, durchaus berücksichtigt worden. Diese detaillierten Überlegungen zur Idee der Cell Assemblies möchten wir im folgenden kurz zusammenfassen und in den Kontext der vereinfachten neuronalen Simulation kognitiver Leistungen stellen.

Der zentrale Gedanke beim Assembly-Konzept ist, daß sich Assemblies durchaus überlappen können und zwar so, daß praktisch jedes Neuron zu vielen verschiedenen Assemblies gehören kann. Das bedeutet für die Gruppierung von Neuronen zu größeren Einheiten oder "Units", daß diese Gruppen in unterschiedlichen funktionellen Kontexten jeweils verschieden gebildet werden können, also nicht von vornherein feststehen. Dadurch erhält man insgesamt eine weit größere Zahl potentieller "Units" als es Kolumnen oder sogar Neurone im Gehirn gibt. Außerdem wird durch eine solche funktionsabhängige Organisation der "Units" eine größere Flexibilität bei der Informationsverarbeitung erreicht, die zumindest theoretisch auch die Ressourcen der vorhandenen Neuronen nahezu optimal nutzt (Palm, 1980, 1982, 1990). Um trotz der Überlappung von Assemblies diese bei der Aktivierung noch gut auseinanderhalten zu können, ist eine starke interne Verkopplung der Neurone innerhalb jeder Assembly nötig. Zu der Frage, wieviele Assemblies mit wie großer Überlappungen auf diese Weise mit insgesamt N Neuronen noch verlässlich realisierbar sind, wurden im Kontext der Theorie der Assoziativspeicher theoretische Überlegungen, konkrete Berechnungen und Computersimulationen angestellt (Palm, 1980; 1991; Palm und Sommer, 1996). Durch die Möglichkeit des Überlapps, der u.U. sogar so weit gehen kann, daß ein Assembly ein anderes vollständig umfaßt, sind zusätzlich zu den einfachen Verbindungsmöglichkeiten noch weitere interessante Beziehungen zwischen zwei "Units" (also Assemblies) in einem (Makro-)Netzwerk möglich. Abbildung 1 zeigt einige Möglichkeiten. Aus dieser Sicht bieten sich die Assemblies als eine geeignete "Zwischenebene" zur vereinfachten Modellierung kognitiver Leistungen an, die zwischen der Ebene der Einzelneurone und der Ebene ganzer Hirnregionen (die heute in fMRI-Untersuchungen beobachtet wird) sinnvoll eingezogen werden kann.

Unserer Vorstellung nach kann ein Assembly aufgefaßt werden als eine über ein gegebenes neuronales Netzwerk verteilte Repräsentation einer "kognitiven Einheit". Im einfachsten Fall mag dies z.B. die Repräsentation eines diskreten externen Reizes oder eines Objektes sein. Ebenso wie jedoch Objekte in der Regel in feinere Bestandteile und Merkmale zerlegt werden können, so können auch Assemblies aus Untereinheiten oder Subassemblies zusammengesetzt sein (Kompositionalität). Zusammengehörigkeit und Relationen zwischen Teilen spiegeln sich hierbei in den Konnektivitäten zwischen den beteiligten Neuronen wider. Dies schließt ein, daß Objektrepräsentationen wiederum selbst zu größeren Strukturen über kontextuelle Zusammenhänge zusammengebunden werden können. Der Begriff des Assemblies wird zudem von uns auch nicht nur auf Repräsentation

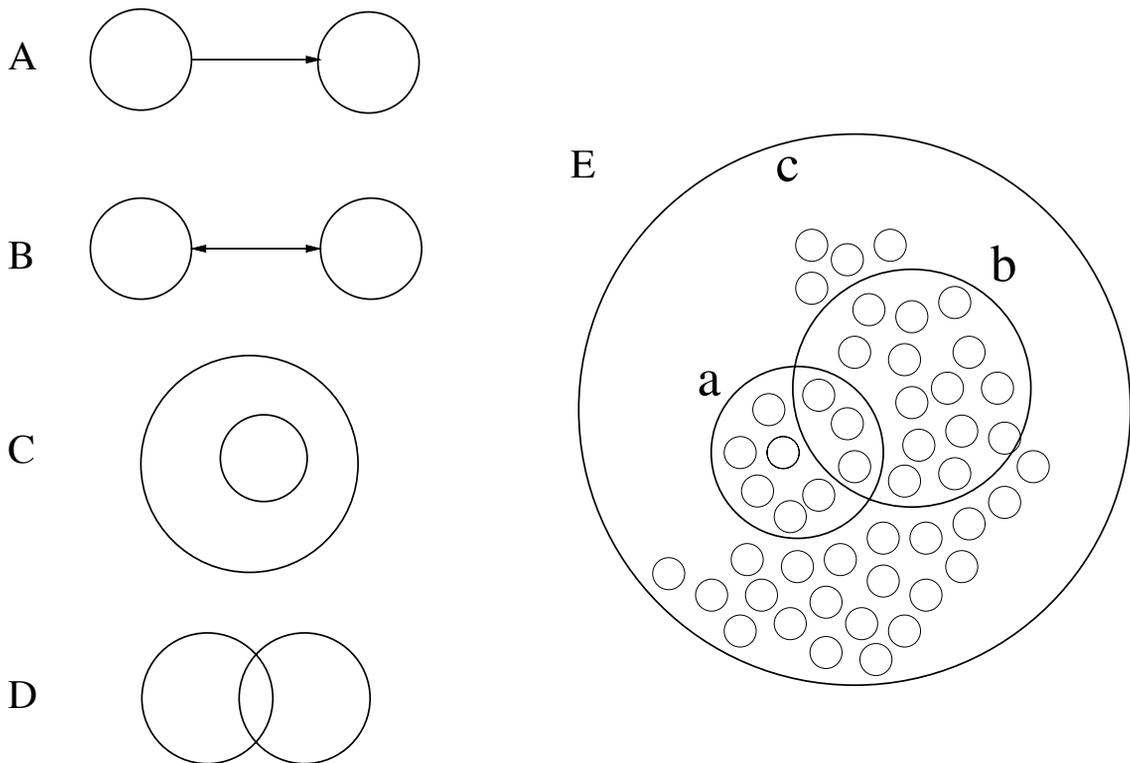


Abbildung 1: Darstellung einiger Relationen zwischen Assemblies. A und B stellen Varianten dar, wie sie auch in klassischen konnektionistischen Modellen möglich sind. In A projiziert das linke Assembly unidirektional auf das rechte und kann dieses bei geeigneter Schwelle assoziativ anregen (Aktivierung). In B ist das aufgrund bijektiver Verbindungen in beide Richtungen möglich (wechselseitige Aktivierung). D unterscheidet sich von B dadurch, daß beide Assemblies gemeinsame Neurone haben (Überlapp). Wieder ist eine wechselseitige Aktivierung möglich; durch Anregung der gemeinsamen Zellen aber auch eine gleichzeitige Adressierung beider Assemblies. In C und E enthält ein größeres Assembly kleinere Subassemblies vollständig (Enthaltensein). Die Subassemblies sind durch relativ zueinander abgestufte Verbindungsstärken definiert. Bei Schwellenvariation können hierdurch Subassemblies (a oder b oder beide) gezielt ausgewählt werden (sie haben intern starke und wechselseitig etwas schwächere Verbindungen zwischen den jeweiligen Neuronen) oder das gesamte Assembly (c) zum Feuern gebracht werden (Verbindungen von a und b nach c sind noch einmal schwächer, aber verglichen mit Verbindungen zu beliebigen anderen Zellen noch stark).

tionen externer Dinge angewendet, sondern in gleicher Weise auch auf “interne” kognitive Entitäten wie Ideen oder Konzepte. Außerdem sollte klar sein, daß das Assemblykonzept nicht direkt im Zusammenhang mit bewußter Wahrnehmung steht. Zwar mögen bewußt wahrgenommene Dinge im Gehirn stets durch momentan aktive Assemblies repräsentiert sein; sicherlich gelangt aber nicht jedes momentane Assembly auch ins Bewußtsein.

Die große Aufgabe, kognitive Leistungen auf der Basis von neuronalen Aktivitätsausbreitungsprozessen zu erklären, zerfällt damit in zwei Teilaufgaben: Nämlich erstens, wie bisher bereits diskutiert, kognitive Leistungen und Funktionen auf der Basis künstlicher neuronaler Netze oder konnektionistischer Modelle zu erklären, wobei die “Units” den Assemblies entsprechen und durch die Überlappung noch weitere interessante Beziehungen zwischen “Units” modellierbar sind, die überdies etwas mit der Darstellung von Ähnlichkeit in den häufig geforderten “verteilten Repräsentationen” zu tun haben. Arbeiten, die diesen Ansatz verfolgen sind u.a. Wickelgren (1969, 1977, 1981, 1992), Braitenberg (1978), Palm (1982, 1990), Damasio (1989), Braitenberg und Pulvermüller (1992), Pulvermüller (1992, 2000), Bibbig et al. (1995, 1996ab), Pasemann (1995), Bienenstock (1995), Wenekers und Palm (1997, 2000).

Zweitens geht es dann darum, die Beziehung zwischen der Aktivierung einzelner Neurone in größeren Netzwerken wie der Großhirnrinde und der Aktivierung von Assemblies herzustellen. Auf diesem Weg ist es zunächst einfacher, die Debatte zu umgehen, inwieweit Assemblies nicht nur lokal sondern sogar global über das ganze Gehirn zusammenhängen können, indem man konzeptuell die vielleicht möglichen globalen Assemblies in lokale Untergruppen zerlegt. Dann stellen sich die folgenden Fragen:

1. Wie stark sind die Neurone eines lokalen Assemblies untereinander verbunden?
2. Wieviele lokale Assemblies kann es in einer Cortexkolumne geben?
3. Wie sieht die Aktivierung eines (lokalen) Assemblies in den zugehörigen Neuronen aus?
4. Wie sieht die Aktivierung eines (lokalen) Assemblies in größeren elektrischen Signalen aus, etwa in lokalen Feldpotentialen oder dem Oberflächen-EEG?
5. Können Assemblies durch Lernen entstehen und wie ist Hebb’sches Lernen in den synaptischen Mechanismen realisiert?

Diese Fragen wurden in weiten Bereichen der experimentellen und der theoretischen Neurobiologie untersucht (Perkel et al., 1967; Gerstein et al., 1989; Singer and Gray, 1995; Palm und Aertsen, 1996; Eckhorn, 1999; Abeles, 1991; Edelman, 1992; Aertsen, 1993; Aertsen et al., 1989, 1994ab; Palm, 1982; Lansner und Fransen, 1992; Koch und Davis, 1994; Arbib, 1995; u.a.).

Die erste Frage hängt eng damit zusammen, wie Teile eines Assemblies relational zusammenhängen (vgl. Abbildung 1). Wenn Subgruppen von Zellen in einem Assembly

Teilaspekte oder Merkmale des repräsentierten Objekts darstellen (oder auch seine Zugehörigkeit zu umfassenderen Kategorien), dann sollten diese Subgruppen entsprechend der Determiniertheit, Sicherheit oder Plausibilität der entsprechenden Relationen unterschiedlich stark miteinander verbunden sein. Korrelationsbasierte Lernregeln, wie Hebb'sche Plastizität, unterstützen dies, indem Zellgruppen, die häufig gleichzeitig aktiv sind, auch stärker miteinander synaptisch verbunden werden (Hebb, 1949; Palm, 1991). Über eine kollektive Modulation neuronaler Feuerschwellen können dann beispielsweise Teile des Assemblies, die nur indirekt mit dem repräsentierten Ding zu tun haben oder selten in Kombination mit ihm auftreten, selektiv ausgeblendet werden; erst bei niedrigen Schwellen werden sie hinzu assoziiert (vgl. Braitenberg, 1978; Palm, 1982).

Die zweite Frage kann im Rahmen von Assoziativspeichermodellen unter Berücksichtigung biologischer Bedingungen mathematisch untersucht werden. Hierzu gibt es eine Reihe von Studien, auf die wir aber nicht im Detail eingehen wollen (eine Übersicht vermitteln Amit, 1988; Palm, 1986; 1991; Palm und Sommer, 1996), sondern als wesentliches Resultat nur erwähnen, daß prinzipiell eine sehr große Zahl von Assemblies in einer Kolumne möglich ist, die sogar die Anzahl der Neuronen in der Kolumne bei weitem übersteigen kann. Dieses theoretische Ergebnis kann allerdings mit den momentan existierenden experimentellen Methoden noch nicht physiologisch überprüft werden.

Probleme des Lernens und der synaptischen Plastizität (Frage 5) möchten wir im weiteren nur am Rande erörtern. Aus neurobiologisch-physiologischer Sicht gibt es hierzu in letzter Zeit viele experimentelle Daten und einige gute Reviews (Tsumoto, 1992; Bliss und Collingridge, 1993; Artola und Singer, 1993; Markram und Tsodyks, 1996; Magee und Johnston, 1997; Zhang et al, 1998). Die Fragen, die wir im folgenden näher untersuchen wollen, sind die Fragen 3 und 4. Außerdem möchten wir beschreiben, wie man den Weg nach oben, also zu beliebig komplexen kognitiven Leistungen auf der Basis der Assemblytheorie zumindest im Prinzip realisieren kann.

Hierzu werden wir von kleinen, im Cortex relativ lokalisierten Netzwerken (z.B. Kolumnen) ausgehen und ihre biophysikalischen Eigenschaften betrachten insofern sie für die obigen Fragestellungen interessant erscheinen. Anschließend werden wir den Zusammenhang zwischen Assemblies und Assoziativspeichern klarstellen und erläutern, wie sich Kurz- und Langzeitgedächtnis in einfacher und schematischer Weise im Assemblykonzept wiederfinden. Schrittweise gehen wir dann zu komplexeren Netzwerken über, indem zunächst gezeigt wird, wie neben statischen Objektbeziehungen auch zeitliche Eigenschaften im Assoziativspeicher repräsentiert werden können. Zudem betrachten wir "globale Assemblies", d.h. solche, die über größere Bereiche, z.B. mehrere Areale des Cortex verteilt sind. In diesem Zusammenhang kommen wir speziell auf das "Bindungsproblem" zu sprechen, sowie auf einige frühere Hirntheorien im engen Zusammenhang mit den vorgestellten Konzepten. Schließlich diskutieren wir, wie diese früheren Modelle um "operationelle" oder regelhafte Komponenten erweitert werden können, die für eine Modellierung kognitiver Leistungen, die über rein assoziative Ansätze hinausgeht, notwendig erscheinen.

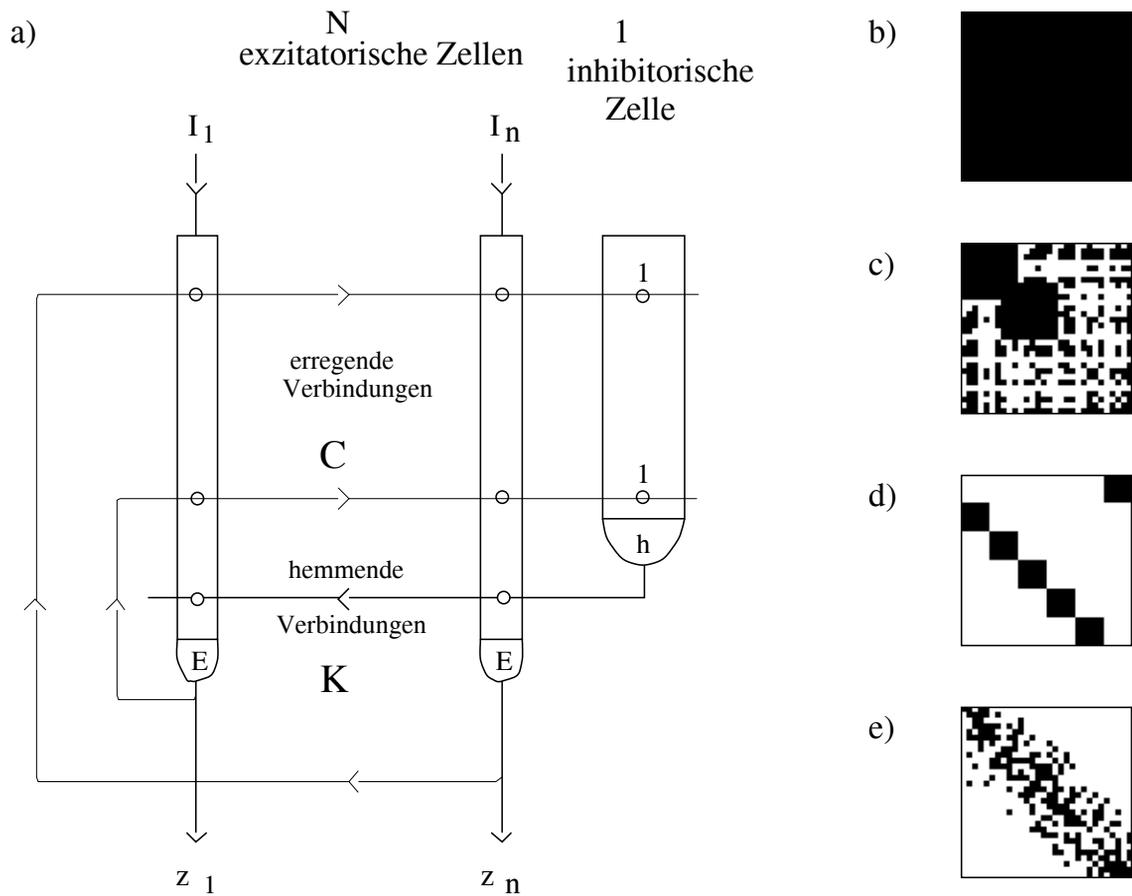


Abbildung 2: Rückgekoppeltes Netzwerk aus N exzitatorischen Neuronen und einer globalen inhibitorischen Zelle, die zur Regulation der mittleren Feuerrate dient und repräsentativ für einen Pool hemmender Zellen steht. Abbildungen b bis e zeigen einfache mögliche Fälle für die Wahl der exzitatorischen Kopplungsmatrix C : b) voll vernetzt (nur ein Assembly), c) Assoziativspeicher, d) Sequenzspeicher, e) zufällige Nachbarschaftskopplungen.

2 Neuronenmodelle und elementare neuronale Architekturen

Ausgehen wollen wir von dem einfachen neuronalen Netz in Abbildung 2 links. Dargestellt sind N exzitatorische Zellen, die über erregende Verbindungen auf sich selbst zurückgekoppelt sind (d.h. massive Divergenz, Konvergenz und Rückkoppelungen sind möglich). Zusätzlich gibt es eine inhibitorische Zelle, die von den exzitatorischen Zellen erregt wird und diese umgekehrt hemmt. Diese Zelle reguliert die Aktivität der exzitatorischen Zellen durch eine negative Rückkoppelungsschleife und verhindert ein unkontrolliertes Anwachsen der Netzwerkaktivität. Anstelle einer einzigen hemmenden Zelle könnte man hier auch mehrere vorsehen, z.B. im Verhältnis 4:1 zwischen exzitatorischen und inhibitorischen Zellen, wie es für cortikale Areale von Säugetieren ein typischer Wert wäre. Unsere Zelle soll jedoch einen ganzen Pool lokaler hemmender Zellen repräsentieren, die sich alle ähnlich verhalten. Dieser Ansatz ist angemessen, insofern hemmende Zellen sehr wahrscheinlich nicht so heterogene Verbindungen durch synaptische Plastizität (Hebbsches Lernen) aufbauen, wie exzitatorische Neurone. Ihr Verhalten ist daher auch homogener und es wird häufig vermutet, daß hemmende Zellen für die Informationsverarbeitung im Cortex nur unterstützende Funktion haben, etwa im Sinne der oben genannten Aktivitäts- oder Schwellenkontrolle. Die "Hauptarbeit" leistet das Netzwerk der exzitatorischen Zellen (Braitenberg und Schüz, 1991).

Zur weitergehenden Modellierung biophysikalischer oder kognitiver Leistungen in dem angegebenen rekurrenten Netzwerkschema müssen zwei wesentliche Dinge festgelegt werden: der genaue Typ der Einzelzellen und die Architektur des Netzwerks, ausgedrückt durch die Kopplungsmatrix C zwischen den exzitatorischen Zellen. Auf der rechten Seite in Abbildung 2 sind einige einfache Kopplungsmatrizen für Architekturen dargestellt, mit deren Hilfe sich jeweils unterschiedliche Gesichtspunkte des Verhaltens neuronaler Netze studieren lassen. Es ist hier angenommen, daß das Netzwerk lediglich 30 Zellen hat; die Kopplungsmatrix hat demnach 30 mal 30 Einträge, nämlich je einen für jede Synapse von einem präsynaptischen Neuron (Zeilen der Matrix, entsprechend den horizontal laufenden axonalen Fasern im Bild links) zu einem postsynaptischen Neuron (Spalten, entsprechend der vertikalen Zellkörper). Ein schwarzer Eintrag in den dargestellten Matrizen bedeutet, daß eine Synapse vorhanden ist, ein weißer, daß keine Verbindung besteht.

Abbildung 2 b zeigt demnach ein Netzwerk, bei dem jede Zelle mit jeder verbunden ist. Dies entspricht z.B. einem einzigen im Netz vorhandenen Assembly. Nach Festlegung eines geeigneten Neuronentyps kann man an einem solchen Netzwerk das qualitative dynamische Verhalten eines Assemblies untersuchen. Hierauf gehen wir weiter unten noch genauer ein.

Abbildung 2 c zeigt eine Assoziativspeichermatrix. Hier sind vier Muster (oder Assemblies) gespeichert. Das erste Assembly besteht aus den ersten 10 Neuronen des Netzes, daher sind die Synapsen zwischen diesen Neuronen alle stark. Das zweite Muster besteht aus den Neuronen 6 bis 15, seine internen Verbindungen erkennt man als zweiten schwarzen Block in der Matrix. Man beachte auch, daß beide Assemblies überlappen, denn sie

haben die Neurone 6 bis 10 gemeinsam. Die beiden anderen eingespeicherten Assemblies bestehen schließlich aus je 10 zufällig ausgewählten Neuronen. Die entsprechenden Synapsen ergeben die weiteren zufälligen Schwärzungen in der Matrix. In diesem einfachen Assoziativspeicherbild sind "Muster" also nichts weiter als Mengen von Neuronen. Sie haben keine direkte semantische Bedeutung. Erst in konkreten Anwendungen (oder im Gehirn) bekommen sie eine Bedeutung indem sie ein externes oder internes kognitives Ereignis repräsentieren, beispielsweise einen Frosch oder einen Scherz.

In Abbildung 2 d sind in dem Netzwerk 6 Muster enthalten, erkennbar an den 6 synaptischen Blöcken. Die Muster bestehen aus je 5 aufeinander folgenden Neuronen, sie überlappen also nicht. Dies aus Gründen der Darstellung: betrachtet man die Abbildung genauer fällt nämlich auf, daß jedes Muster nur Synapsen zum jeweils nächsten Muster hat. Daß heißt die Aktivierung eines Musters führt zur jeweiligen Aktivierung des Folgemusters. Es kommt so zu einer gerichteten Ausbreitung der Aktivität im Netzwerk, die hier zyklisch ist, weil die 6 Muster zu einem Ring zusammengeschlossen sind. Solche Netzwerke können für Modelle raum-zeitlicher Anregungen im Gehirn dienen, für die es auch neurophysiologische Evidenz gibt (sog. synfire chains, Abeles et al., 1991; Abeles et al., 1993). Letztendlich sollte auch Gedächtnisinhalten mit zeitlicher Dimension eine derartig gerichtete Aktivitätsausbreitung in cortikalen Netzwerken entsprechen. Die Architektur in 2 d ist also ein einfaches Beispiel dafür, wie "Zeit" in Modelle neuronaler Netze eingeführt werden kann. Auch hierzu später mehr.

Abbildung 2 e zeigt schließlich ein Beispiel für "topographische" Kopplungen. Die Wahrscheinlichkeit für eine Synapse zwischen zwei Neuronen ist hier um so geringer, je weiter die beiden Zellen auseinander liegen. Solche Netze kann man auch leicht zweidimensional formulieren und etwa als Modell für ein ausgedehntes Cortexareal auffassen. Man kann an ihnen dann unter anderem die laterale Aktivierungsausbreitung in Cortexarealen studieren und versuchen derartig simulierte Signale in Zusammenhang mit Wellenphänomenen zu bringen, wie sie in EEG-Ableitungen gefunden werden (Wright und Liley, 1996; Nunez, 1995). Erweiterte Modelle dieser Art zur Beschreibung primärer sensorischer Areale berücksichtigen zudem noch rezeptive Feldeigenschaften in den Einzelzellen, oft in Form von Karten für die entsprechenden Merkmalsklassen (z.B. Orientierungs- und Richtungsselektivität in Modellen für das visuelle System). Mithilfe solcher detaillierter Modelle läßt sich hervorragend die Informationsverarbeitung in sensorischen (z.T. auch motorischen) primären Arealen untersuchen (z.B. Wörgötter und Koch, 1991; Erwin et al., 1995; Stöcker et al., 1996; Kirkland und Gerstein, 2000). Da diese jedoch in erster Linie in Vorverarbeitungsprozessen und speziell in einer immensen Reduktion der einlaufenden Sinnesdaten besteht - was beides als noch wenig kognitiv angesehen werden kann - werden wir im weiteren solche Modelle nicht diskutieren.

Wie schon gesagt gehört zur vollständigen Spezifikation eines neuronalen Netzes neben seiner Architektur auch die Wahl eines geeigneten Neuronenmodells. Hier hat man eine große Auswahl auf ganz unterschiedlichen Niveaus biologischer Detailtreue. Beschränkender Faktor ist zum einen die zur Verfügung stehende Rechnerleistung, zum andern aber

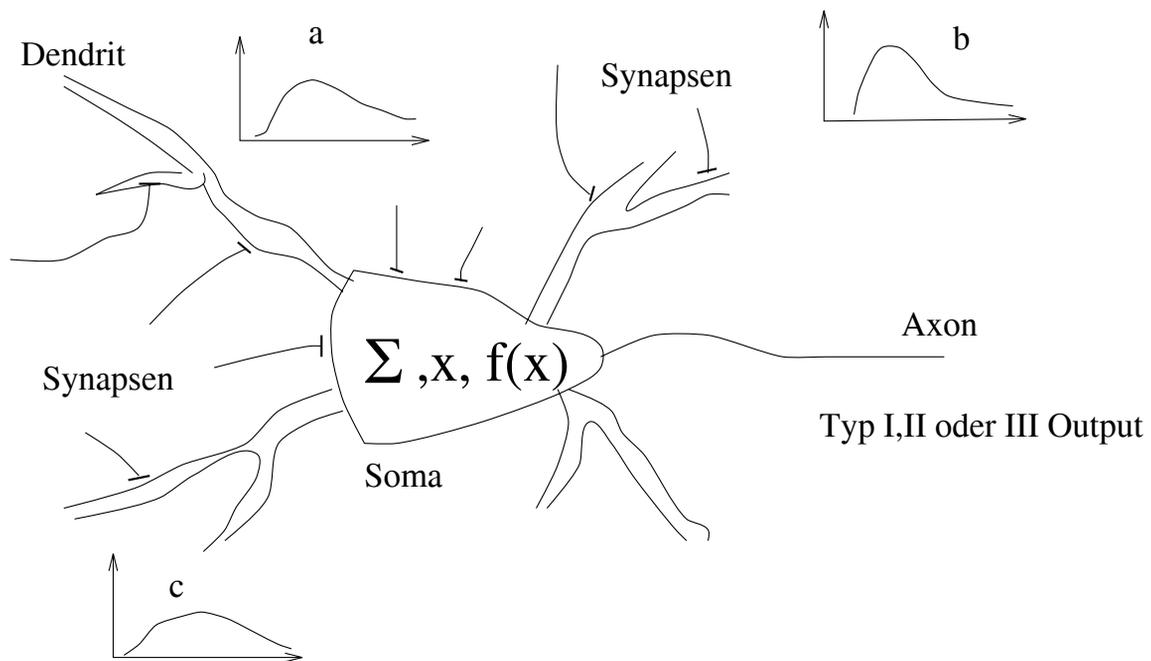


Abbildung 3: Einzelneuronschema: Das Neuron wird von anderen Zellen über Synapsen auf seinen Dendriten und dem Soma erregt. Einzelne Synapsen können als Antwort auf pulsformige Aktionspotentiale spezifischen Spannungsänderungen der Zellmembran hervorrufen (a, b, c). Durch Summation (Σ) wird auf dem Soma das Generatorpotential x berechnet, aus dem sich die mittlere Feuerrate der Zelle $f(x)$ ergibt, die entweder als gradueller Wert (Typ I) oder in Form von Spikes (Typ II/III) über das Axon ausgegeben wird.

natürlich auch das konkret betrachtete Problem, denn viele Phänomene, die in Netzen mit kompliziert aufgebauten Einzelzellen auftreten, findet man oft in qualitativ ganz ähnlicher Weise, wenn man statt dessen ein einfacheres Neuronenmodell verwendet. In dem Fall steht die Verwendung des komplizierten Modells der Klarheit der Darstellung sogar oft im Wege. Das gilt speziell auch für Phänomene, die sich als Netzwerkeffekte ergeben, also daraus, daß eine große Zahl von Zellen miteinander gekoppelt sind und wechselwirken. Details und kleine Unterschiede in den Einzelneuronen verschwinden dann leicht in dem kollektiv zu beobachteten Verhalten des Netzwerks, sie mitteln sich im kollektiven Verhalten heraus. Gelegentlich ist es sogar möglich, eine solche Vereinfachung mathematisch zu begründen (Harth et al., 1970; Little, 1974; Amari et al., 1977). Eine solche Mittelung bildet in gewisser Weise die Grundlage dafür, daß überhaupt kognitive Modelle mit größeren "Einheiten" als es die Einzelneuronen sind, sinnvoll formuliert und untersucht werden können. Hierfür geben wir weiter unten noch ein Beispiel an.

Modelle für Einzelzellen bestehen grob gesehen aus zwei Komponenten: ein Teil beschreibt die synaptische Summation einlaufender Signale auf den Dendriten und dem Soma der Zelle, der zweite Teil beschreibt die Generierung von Ausgabesignalen aus dem

dendritisch berechneten Summenpotential. Diese Ausgabesignale werden dann wiederum an andere Zellen weitergegeben.

Die Berechnung des Summenpotentials erfolgt in den einfachsten Modellen als eine gewichtete Summe über die Eingangssignale, wobei die Gewichte den synaptischen Stärken wirklicher Neurone entsprechen. Berücksichtigt man Lernprozesse im Netzwerk auf der Grundlage synaptischer Plastizität, sind diese Gewichte zudem noch gemäß einer vorgegebenen, z.B. Hebb'schen, Lernregel veränderlich. In etwas aufwendigeren Modellen kann man zusätzlich noch das zeitliche Verhalten von synaptisch hervorgerufenen Potentialschwankungen auf den Zellmembranen genauer beschreiben. Dies geschieht in Form von Differentialgleichungen und ist notwendig, wenn man kognitive Phänomene studiert, bei denen es speziell auf das Zeitverhalten von Neuronengruppen ankommt. Ein Beispiel hierfür ist das sog. "Binden durch Synchronisation", gemäß dem Neuronen ihre Zugehörigkeit zur Repräsentation desselben Objekts durch synchronisiertes rhythmisches Feuern im Gammafrequenzbereich (30-90Hz) ausdrücken (Eckhorn et al., 1988; Eckhorn, 1999; Gray, 1994; Singer und Gray, 1995). Solche Phänomene kann man mit simplen zeitdiskreten konnektionistischen Ansätzen nicht erfassen.

Die einfachste Form der Ausgabeerzeugung in Neuronenmodellen besteht darin, mittels einer nichtlinearen Funktion aus dem Summenpotential einer Zelle eine Rate zu berechnen, die angibt mit welcher mittleren Frequenz die entsprechende Zelle bei diesem Potential Aktionspotentiale generieren (d.h. "feuern") würde. Aktionspotentiale, auch "Spikes" genannt, werden aber selbst nicht simuliert, sondern es wird der Ratenwert direkt an die anderen Zellen weitergegeben. Dies ist die übliche Vorgehensweise in konnektionistischen Modellen. Ein solches "Neuron" muß man als Pool von Zellen interpretieren, die asynchron mit ungefähr der bestimmten Rate feuern. Der kollektive Einfluß aller Neurone im Pool auf eine Zielzelle ist dann in guter Näherung nur durch die Feuerraten beschreibbar.

In den letzten Jahren hat sich jedoch zunehmend gezeigt, daß einfache Ratenmodelle zur Beschreibung und Erklärung von Informationsverarbeitungsprozessen in neuronalen Netzen nicht immer ausreichen. Es gibt wachsende Evidenz dafür, daß Information nicht nur in der mittleren Rate eines Neurons, sondern auch in seinen Feuerzeitpunkten kodiert ist (Eckhorn et al., 1988; Abeles, 1991; Singer und Gray, 1995). Dies trifft beispielsweise auf das schon genannte "Binden durch Synchronisation" zu, wo das gleichzeitige Feuern von Zellen Zusammengehörigkeit kodieren soll, Objektteile zu Objektrepräsentationen bindet und hierdurch kognitiv relevant wird. In solchen Fällen kommt man nicht umhin, Aktionspotentiale oder Spikes wenigstens in einer einfachen Form in der Modellierung zu berücksichtigen.

Die einfachste Form dies zu tun ist, von der genauen Form von Aktionspotentialen abzusehen und Spikes als binäre 0/1-Folgen zu modellieren. Die Ausgaben der Modellneurone sind dann nicht mehr Raten, sondern Folgen von Einsen und Nullen. Eine Eins bedeutet, daß die Zelle zum jeweiligen Zeitpunkt feuert, eine Null, daß sie inaktiv ist. Auf diese Weise können Synchronisationseffekte und andere Korrelationen zwischen Neuronen sinnvoll beschrieben werden.

Die Ausgabespikes können auf verschiedene Weise aus dem dendritischen Summenpotential bestimmt werden. Eine oft verwendete Möglichkeit ist, die wie zuvor mittels nichtlinearer Ausgabefunktion bestimmte Rate als Wahrscheinlichkeit zu interpretieren, mit der zu einer gegebenen Zeit ein Spike erzeugt wird. Die Spikeerzeugung ist dann rein probabilistisch und einzelne Neuronen generieren sogenannte "ratenmodulierte oder inhomogene Poissonprozesse".

Etwas aufwendiger können Spikefolgen auch mit Hilfe einer zellintrinsic dynamischen Feuerschwelle erzeugt werden. Hier wird jedesmal wenn das Summenpotential die Schwelle überschreitet ein Spike generiert. Gleichzeitig wird die Schwelle für eine gewisse Zeit erhöht, so daß die Zelle nicht sofort wieder feuert, sondern eine Weile "refraktär" ist. Natürlich kann man diese deterministische Spikeerzeugung auch mit der probabilistischen kombinieren, etwa indem man noch Rauschen auf der Membran oder in der Schwelle berücksichtigt. Solche einfachen spikenden Neurone beschreiben elementare Eigenschaften des neuronalen Feuerverhaltens schon recht gut (Perkel et al., 1967; Stein, 1967; Lankheet et al., 1989; Gerstner, 1995).

Gelegentlich ist es notwendig, noch über diese einfachen Modelle spikender Neuronen hinaus zu gehen. Dann kommt man zu "Kompartimentmodellen" die sowohl die Anatomie als auch die Physiologie wirklicher Neuronen abzubilden versuchen. Zum einen wird die Form einer Zelle durch eine Anzahl Kompartments nachgebildet (2-2000), zum anderen wird auf jedem Kompartment das elektrische Verhalten der Membran unter Berücksichtigung der wichtigsten beteiligten Ionenkanäle und Neurotransmitter genau nachgebildet (Koch und Segev, 1989; Traub, 1991; Pinsky und Rinzel, 1994). Die Verwendung solcher Modelle läßt momentan noch kaum die Simulation größerer Netzwerke zu, was sich aber in Zukunft mit der Bereitstellung schnellerer Computer ändern wird.

Auch wenn eine möglichst einfache Modellierung aus theoretischer Sicht wünschenswert ist, sollte man trotzdem stets im Auge behalten, daß sich manchmal auch Details in der Funktionsweise von Einzelzellen in kognitiven Leistungen widerspiegeln können (Segev, 1992). Dies wird etwa offensichtlich, wenn man sich den Einfluß gewisser Pharmaka auf die Ausführung kognitiver Aufgaben vor Augen führt. Hier spielen Prozesse auf der Neurotransmitterebene eine Rolle, die man in zu einfachen konnektionistischen Modellen nicht adäquat beschreiben kann. Ganz ähnlich wirken neuromodulatorische Systeme auch unterschiedlich in verschiedenen Verhaltenssituationen (Wachheit, Aufmerksamkeit, Erregung, Furcht, etc.) wo sie zweifellos Einfluß auf konkrete Handlungen, neuronale Aktivierungsmuster (z.B. auch im EEG) und sogar synaptische Plastizität haben. Solche Phänomene wird man letztendlich kaum verstehen können ohne eine angemessenen Modellierung auf relativ niedriger Ebene. Für unsere Zwecke reichen aber zunächst einfachere Ansätze, da es uns überhaupt erst einmal um einen groben Entwurf davon geht, wie im Gehirn kognitive Phänomene auf der Grundlage von Assemblies von statten gehen könnten.

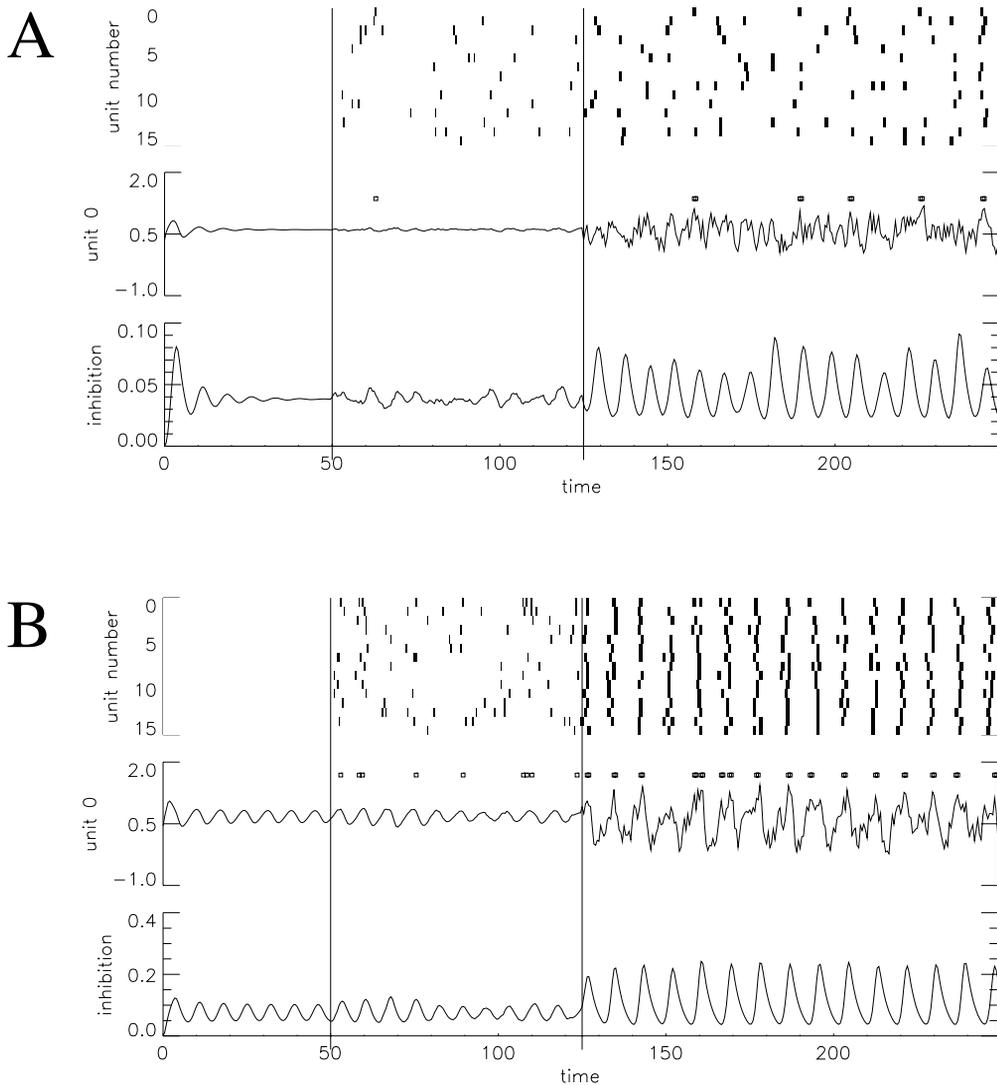


Abbildung 4: Qualitatives Verhalten der drei im Text beschriebenen Einzelneuronvarianten in einem vollständig gekoppelten exzitatorischen Netzwerk mit inhibitorischer Aktivitätskontrolle (vgl. Abb. 2 a und b). In A und B werden während der ersten 50 Zeiteinheiten Neuronen vom graduellen Typ (I) benutzt; von 50-150 stochastische Schwellenneurone (II) und danach solche mit Refraktärmechanismus (III). Der obere Plot zeigt in A und B jeweils die Spikes (schwarze Punkte) von 16 der insgesamt 1024 simulierten Zellen; der mittlere Plot gibt exemplarisch Membranpotentiale und Spikes (Punkte) einer Einzelzelle wieder, und unten ist jeweils der Zeitverlauf des inhibitorischen Neurons zu sehen. In Bild A ist die Aktivität von Typ I und II Neuronen (abgesehen von Fluktuationen des Membranpotentials) stationär, wohingegen sie bei Verwendung von Refraktärneuronen oszillatorisch ist. Bei höherer externer Anregung wie in Abbildung B beginnen auch die Netzwerke aus Typ I und II Zellen zu schwingen. Diese Oszillationen liegen bei Verwendung plausibler Parameterwerte im EEG-Gammawellenbereich.

3 Lokale Netzwerke und Zeitverhalten von Assemblies

Zunächst betrachten wir einmal das qualitative Zeitverhalten eines einzelnen Assemblies modelliert wie in Abbildung 2 a mit der vollständigen Kopplungsmatrix in Abbildung 2 b. Zum Vergleich enthält jede der Abbildungen 4 A und B Simulationen des Zeitverlaufs dreier Modellvarianten mit jeweils verschiedenen Einzelzellmodellen: In den ersten 50 Zeiteinheiten (entsprechend etwa 100-150 ms "Echtzeit") werden je 1024 graduelle exzitatorische Zellen benutzt (Typ I), von 50 - 125 solche mit probabilistischer Spikegenerierung (Typ II) und anschließend Neurone mit Refraktärmechanismus (Typ III). In diesen Intervallen ändert sich jeweils nur die Generierung der Ausgangssignale und zwischen beiden Bildern nur die Intensität der externen Anregung; die Netzwerkstruktur, Membrandynamik und Parameter bleiben stets dieselben. Der obere Plot in A und B zeigt nun jeweils ein zeitliches Raster der Feuerzeitpunkte von 16 der insgesamt 1024 Zellen. Da graduelle Zellen keine Spikes generieren, fehlen sie im Zeitbereich 0-50. Der mittlere Plot enthält exemplarisch Membranpotential und Feuerzeiten (Punkte) einer einzelnen Zelle, während unten das Potential der hemmenden Zelle dargestellt ist, das in etwa die mittlere Aktivität der exzitatorischen Zellen widerspiegelt, denn die hemmende Zelle summiert ja über die Ausgangsaktivität der exzitatorischen Neurone. Zur Zeit $t=0$ wird die Anregung in Form eines konstanten Eingangsstroms in alle Zellen sprunghaft eingeschaltet. Während der Anfangsphasen der Simulationen zeigen sich daher schnell abklingende Transienten, die man auch in wirklichen neuronalen Netzen findet.

Abbildung 4 soll exemplarisch verdeutlichen, wie Simulationen neuronaler Assemblies auf einer Ebene mittlerer Komplexität aussehen. Einerseits werden physiologische Eigenschaften biologischer Neurone plausibel berücksichtigt, andererseits liegt den Simulationen aber schon der Assemblygedanke als Repräsentationen kognitiver Ereignisse zugrunde.

Weiterhin demonstriert Abbildung 4, daß die Wahl des Neuronenmodells sich auf die Dynamik des Netzwerkverhaltens auswirkt. In A ist das Modell mit Refraktärneuronen (Typ III) nämlich in einem oszillatorischen Bereich, während es für Typ I und II-Neurone stationär ist, wobei sich im Falle von Typ II-Neuronen aber noch Fluktuationen in den Membranpotentialen zeigen, die auf die probabilistische Spikeerzeugung zurückzuführen sind. Die Verwendung dieser beiden Einzelzellvarianten führt erst bei höherer Anregung zu kollektiven Schwingungen, zu sehen in Abbildung 4 B. Dort befindet sich das Netz aus Refraktärneuronen aber schon in einem Anregungszustand, in dem praktisch alle Zellen in jeder Oszillationsperiode feuern. D.h. die verschiedenen Modellvarianten sind nicht äquivalent; eine genauere Betrachtung zeigt, daß sogar die Ursache der Schwingungen für Typ III-Neurone einerseits und Typ I/II-Neurone andererseits unterschiedlich ist. Die gezeigten Schwingungen können als Gammaoszillationen interpretiert werden. Insofern solche das "Binden durch Synchronisation" unterstützen können, kommt es also auf die Wahl des Einzelneuronmodells in Simulationen an, wenn man die entsprechenden Phänomene wirklich verstehen will. Ähnliches trifft im übrigen auch auf kognitive Phänomene in anderen

EEG-Wellenbereichen zu, z.B. auf alpha-, theta- oder delta-Wellen, die in zunehmendem Maße mit kognitiven Phänomenen in Zusammenhang gebracht werden (z.B. Klimesch, 1994; Miller, 2000); zum Teil ist jedoch nicht einmal ihre genaue physiologische Ursache bekannt.

Sieht man von den Ausgangsspiques und dem von ihnen verursachten Rauschen bei den Neuronen vom Typ II ab, zeigen die Zellvarianten I und II im Netzwerk sehr ähnliches Verhalten. Dies sieht man besonders im Potential der hemmenden Zelle in Abbildung 4 A und B unten, das ungefähr die mittlere Aktivität im Netzwerk wiedergibt. Tatsächlich kann man mathematisch nachweisen, daß sich in großen Netzwerken (mindestens einige hundert bis tausend Zellen) beide Netzwerkvarianten praktisch gleich verhalten, und mehr noch, daß das kollektive Verhalten (Stationarität in A, Schwingungen in B) sich genau durch ein reduziertes Netzwerk beschreiben läßt, das zusätzlich zu der einen hemmenden Zelle nun alle exzitatorischen Zellen in einer einzigen, quasi repräsentativen, graduellen Zelle zusammenfaßt (Harth et al., 1970; Amari et al., 1977; Wennekers, 1999). In diesem Sinne kann man das ganze Assembly durch eine reduzierte "Unit" ersetzen, die aus nur einem repräsentativen exzitatorischen und inhibitorischen Neuron besteht. Dies geht ähnlich auch für Netze mit Refraktärneuronen, was z.B. durch das nahezu synchrone Feuerverhalten aller Zellen in Abbildung B nahegelegt wird; das "reduzierte" Modellneuron unterscheidet sich in diesem Fall allerdings von dem für Neuronen vom Typ I und II. Diese Möglichkeit der Dimensionsreduktion rechtfertigt auch die Verwendung nur einer einzigen graduellen hemmenden Zelle in unseren Simulationen (siehe Bild 2 a).

Die Reduzierbarkeit des vollständigen Netzwerks auf repräsentative Units, welche die kollektive Dynamik wiedergeben, bedeutet natürlich, daß sich die Zellen des Netzwerks "ähnlich" verhalten. In Abbildung 4 A feuern die Typ II-Zellen z.B. stationär und mit ähnlicher Rate; in B feuern die Typ III-Zellen näherungsweise einen Spike in jeder Periode. Dieses kollektive Verhalten ist dann letztendlich auch in gemittelten Signalen wiederzuerkennen und durch reduzierte "Units" beschreibbar. Dies spielt eine wichtige Rolle, wenn man nach der physikalischen Beobachtbarkeit von Assemblies in Massensignalen wie lokalen Feldpotentialen, dem Oberflächen-EEG oder neueren bildgebenden Verfahren fragt. Ein Problem ist hierbei die jeweilige räumliche und zeitliche Auflösung des Meßverfahrens, ein zweites, die Größe und "Verteiltheit" des Assemblies. Zeitmuster in kollektiven Signalen, wie z.B. Gammaoszillationen, brauchen für die Beobachtung einerseits natürlich eine entsprechende zeitliche Auflösung der Meßapparatur, andererseits müssen genügend große Gruppen von Zellen kollektiv schwingen, damit das überhaupt in Feldpotentialen oder dem EEG signifikant über dem Hintergrundrauschen nachweisbar wird. Zudem dürfen die Zellen, selbst wenn es einigermaßen viele sind, nicht beliebig über den Cortex verteilt sein, sondern sie sollten möglichst benachbart sein, so daß sie einen kollektiven Einfluß auf den jeweiligen Meßsensor ausüben können. Glücklicherweise wird dies durch die Karten- und Kolumnenstruktur des Cortex unterstützt, nach der Zellen mit ähnlichen Eigenschaften relativ lokalisiert in entsprechenden Arealen zu finden sind.

Bisher haben wir nur das Verhalten eines einzigen Assemblies betrachtet. Der Assozia-

tivspeicher kann als Modell eines neuronalen Netzwerks aufgefaßt werden, in dem mehrere oder viele Assemblies als Muster gespeichert sind, d.h. die entsprechenden Neuronen unterstützen sich gegenseitig durch zahlreiche, verstärkte synaptische Verbindungen. Wie weiter oben schon ausgeführt, kann durch den Überlapp zwischen Speichermustern die Ähnlichkeit der repräsentierten Entitäten ausgedrückt werden, sowie weitere relationale Zusammenhänge zwischen ihnen. Assoziativspeicher sind von uns (und anderen) auf allen oben genannten Ebenen biologischer Realitätsnähe simuliert worden (Willshaw et al., 1969; Palm, 1980, 1982a; Hopfield, 1982; Amit, 1988, 1995; Gerstner und van Hemmen, 1993; Lansner und Fransen, 1992; Renart et al., 1999ab; Sommer und Wennekers, 1998, 2000; Wennekers et al., 1995, 1997, 1999). Ein qualitatives Beispiel zeigt Abbildung 5.

Das für Abbildung 5 verwendete Netzwerk bestand aus 100 exzitatorischen Neuronen und einer hemmenden graduellen Zelle (vgl. Abbildung 2 a). Als Einzelneuronen wurden biologisch realistische Zweikompartmentneuronen verwendet (vgl. Wennekers und Sommer, 1999; Pinsky und Rinzel, 1994). Einfachere spikende Einzelneurone ergeben aber praktisch dieselben Resultate wie im Bild. Gespeichert waren in dem Netzwerk 50 Muster mit jeweils 10 aktiven Neuronen. Das erste Muster bestand aus den Zellen 1-10 und die anderen 49 Muster hatten aktive Neurone an zufälligen Positionen. Im unteren Teil des Bildes sind die Spikes der 100 Zellen über der Zeit als schwarze Markierungen dargestellt, wobei das erste Muster durch konstante Eingangsströme in die ersten 5 Zellen (1-5) adressiert wurde. Die obere Kurve gibt an, wie gut die Netzwerkaktivität (i.e. die Spikes) in den vergangenen 7ms zu dem gerade adressierten Speichermuster paßt. Hierzu wurde ein Informationsmaß verwendet (siehe Wennekers und Sommer, 1999); eine Qualität von 1 entspricht perfekter Vervollständigung.

Man beobachtet in Abbildung 5 eine rhythmische Netzwerkaktivität mit einer Periodenlänge im Gammabereich (ca. 20ms, entsprechend 50Hz). In jeder Periode feuern zunächst die adressierten Zellen 1-5. Nach einer kurzen Verzögerung im Millisekundenbereich, die durch die Rückkopplung der Aktionspotentiale der Zellen 1-5 bedingt ist, feuern zusätzlich die Zellen 6-10, d.h. das erste Muster wird assoziativ komplettiert (quality = 1). Zudem feuern in manchen Perioden auch noch einige andere Zellen, besonders in der ersten Periode gleich nach dem Einschalten des Reizes (quality < 1). Diese zusätzlichen Spikes stellen Vervollständigungsfehler dar, die zu Beginn der Simulation etwas größer sind, da die Zellen im Netzwerk (sowie die Inhibition) in gewissem Rahmen adaptieren. Nach 100 Millisekunden wird vom ersten Muster mit den Neuronen 1-10 auf ein weiteres der zufälligen eingespeicherten Muster umgeschaltet. Wieder werden 5 Zellen dieses neuen Musters adressiert. Die Zellen 1-10 des ersten Musters sind im zweiten Muster nicht enthalten; sie feuern daher nicht mehr im Bereich 100-200ms. Welche Neuronen dieses zweite Muster ausmachen, sieht man in der letzten Gammaperiode etwa zur Zeit 190ms; die 5 am frühesten feuernenden Zellen sind hier wieder die adressierten Zellen und die späteren 5 die hinzu assoziierten, so daß die Verhältnisse während der zweiten Retrievalphase von 100-200ms und der ersten Phase von 0-100ms völlig äquivalent sind, nur daß das adressierte Muster jeweils ein anderes ist. Bemerkenswert ist allerdings, daß das Netzwerk

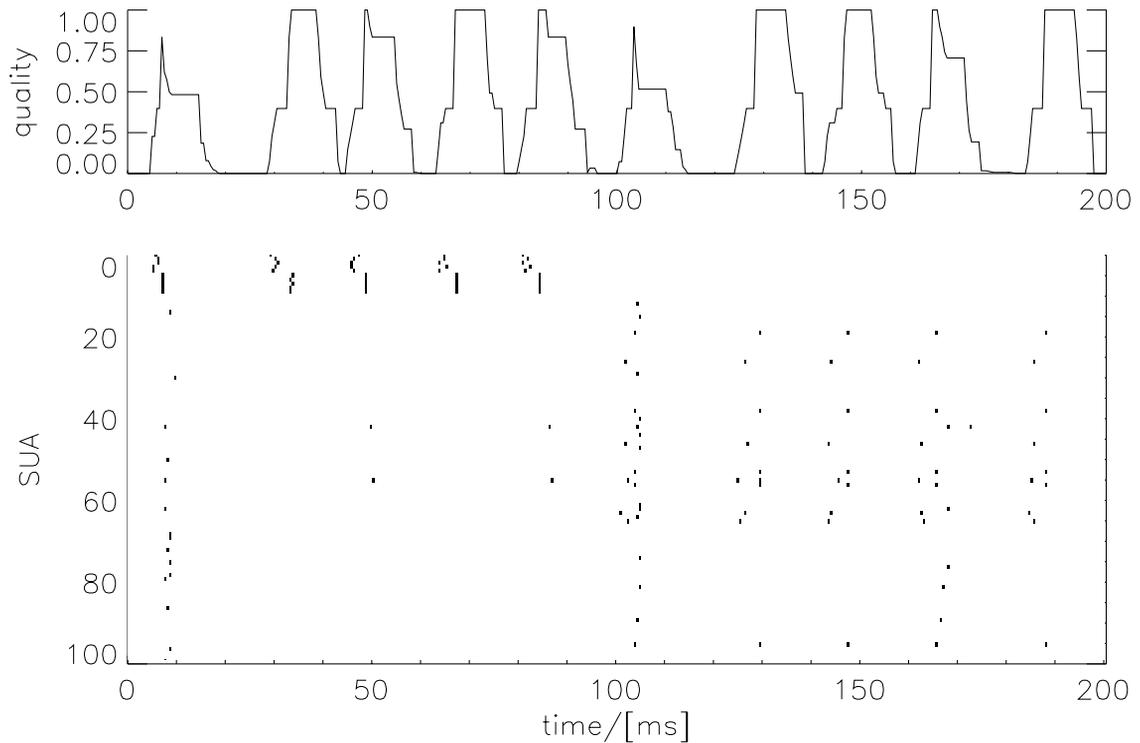


Abbildung 5: Simulation der Mustervervollständigung in einem Assoziativspeicher aus 100 Neuronen in dem 50 Muster mit je 10 aktiven Neuronen abgelegt sind. Das erste Muster besteht aus den Zellen 1-10 von dem in der Zeit von 0-100ms die Zellen 1-5 eine externe Anregung erhalten (sie adressieren das gespeicherte Muster). Die anderen 49 Muster haben je 10 aktive Neurone an zufälligen Positionen. Zur Zeit $t=100$ wird der Input auf 5 Zellen eines dieser Muster umgeschaltet. Der untere Rasterplot zeigt die Spikes aller Zellen über der Zeit, der obere Plot die Qualität des Retrievals ($quality = 1$ entspricht perfekter Vervollständigung). Man beachte, daß die Netzwerkaktivität rhythmisch im Gammabereich ist, sowie daß in jeder Periode zunächst die adressierten Zellen feuern und erst dann die hinzu assoziierten (i.e. 6-10 für Muster 1 während der Zeit von 0 bis 100ms). Gelegentlich feuern auch falsche Neurone, besonders kurz nach dem Umschalten des Stimulus auf ein neues Muster. Hierbei handelt es sich um Adaptationseffekte.

sofort ohne längeres Einschwingen auf das Umschalten des Stimulus zur Zeit $t=100\text{ms}$ reagiert. Mustervervollständigung kann also extrem schnell sein, so schnell, wie es die typischen Membranzeitkonstanten der Neurone im Bereich einiger Millisekunden nur zulassen. Weiterhin ist interessant, daß immerhin 50 Muster mit je 10 aktiven Neuronen in einem Netzwerk aus nur 100 Zellen gespeichert werden. Theoretische Analysen zeigen, daß dies eine bemerkenswert hohe Speicherkapazität darstellt. Das Netzwerk nutzt seine Ressourcen daher nahezu optimal (Palm, 1980; 1990; Schwenker et al., 1996; Wennekers und Sommer, 1999).

Die im Bild 5 gezeigten Oszillationen liegen bei plausibler Wahl der Modellparameter im Gammabereich (ca. 30-90Hz). Sie können daher "das Binden durch Synchronisation" unterstützen. In diesem Zusammenhang sind in den letzten Jahren eine Vielzahl von Modellstudien publiziert worden (Damasio, 1989; Erb und Aertsen, 1992; Tononi et al., 1992; Gerstner und van Hemmen, 1993; Gerstner, 1995; Schillen und König, 1994; Wennekers und Palm, 1997; und viele mehr). Zudem ist es möglich, bei der gleichzeitigen Adressierung mit Teilen zweier oder mehrerer gespeicherter Muster, sowohl eine Vervollständigung der Einzelmuster zu erreichen, als auch eine Segmentierung der Muster, entweder in unterschiedliche Perioden der kollektiven Schwingung oder in verschiedene Phasen innerhalb einer einzigen Schwingungsperiode (Milner, 1974; Gerstner und van Hemmen, 1993; Jensen und Lisman, 1996; Wennekers und Palm, 1997, 1999; u.a.).

In Abbildung 5 erreicht die Netzwerkdynamik (solange der Stimulus sich nicht ändert) nach einer kurzen Einschwingphase mit relativ vielen Vervollständigungsfehlern einen stationären periodischen Netzwerkzustand (limit-cycle) in dem in jeder Periode die Zellen des Speichermusters relativ scharf lokalisiert in der Zeit feuern. Bei anderer Wahl der Netzwerkparameter ist es gleichfalls möglich, daß sich ein Zustand einstellt, in dem die Zellen des gespeicherten Musters asynchron und nicht zeitlich lokalisiert in rhythmischer Weise feuern. In solchen asynchronen Zuständen ist es zudem möglich, daß sich die Aktivität im Netzwerk hält, auch wenn man den Stimulus nach kurzer Zeit wieder abstellt. Die Neurone erregen sich dann nämlich hinreichend stark über ihre wechselseitigen synaptischen Verbindungen, so daß eine zusätzliche externe Anregung überflüssig wird (Hystereseeffekt). Dieser Mechanismus stellt ein einfaches Modell für die Kurzzeitspeicherung von Information in Form rekurrenter Netzwerkaktivität bereit (Harth et al., 1970; Little, 1974; Shaw et al., 1982; Zipser et al., 1993).

Kopplungsstrukturen mit assoziativen Eigenschaften wie in Abbildung 2 c, können in neuronalen Netzen durch Hebbsches Lernen eingespeichert werden. Im einfachsten Fall geschieht dies vor der eigentlich Simulation der Netzwerkdynamik durch Speicherung aller Muster mit Hilfe von auto-assoziativen one-step Lernregeln (Palm, 1982, 1991). Die synaptischen Stärken sind dann während der Simulation fixiert. Biologisch plausibler ist, daß neuronale Netze kontinuierlich lernen, also die Synapsen gemäß Hebbscher (oder anderer) Lernregeln ständig plastisch sind. Auf diese Weise bilden sich neuronale Assemblies in neuronalen Netzen durch Selbstorganisation bei externer Aktivierung durch ein Stimuluster und die hiervon evozierte Feueraktivität der Neurone (Hetherington und Shapiro,

1993; Bibbig et al., 1995, 1996a). Motiviert durch neurophysiologische Befunde im Zusammenhang mit synaptischer Plastizität (long-term potentiation, long-term depression, etc.) werden zunehmend realistischere Lernregeln auf ihre qualitativen Eigenschaften hin untersucht (Abbott und Blum, 1996; Hünig et al., 1998).

Mehrere Dinge machen den Assoziativspeicher interessant für die kognitive Modellierung:

1. Zunächst natürlich die Assoziativität an sich. Hierdurch kann aus bestimmten Schlüsselinformationen auf damit zusammenhängende abgespeicherte Informationen geschlossen werden. Außerdem kann ein Muster aus der Kenntnis einiger seiner Teile oder einer gestörten, verrauschten oder verzerrten Variante vervollständig bzw. wiederhergestellt werden. Wenn das Assembly intern durch relativ zueinander abgestufte Verbindungsstärken noch weiter strukturiert ist, kann zudem über eine Schwellenkontrolle die Größe (i.e. Detailliertheit) des rekonstruierten Assemblies beeinflusst werden (vgl. Abbildung 1 E und auch Braitenbergs Gedankenpumpe (1978)).

2. Sein Kategorisierungsvermögen. Dadurch, daß Varianten des Musters dasselbe Assembly anregen, wird die neuronale Repräsentation in gewissem Grade unabhängig von Feinheiten in der gegenwärtigen Erscheinungsform eines gespeicherten Objekts. Dies mag eine Begründung für das psychophysische Phänomen der Stimuluskonstanz darstellen (Hebb, 1949).

3. Der Assoziativspeicher hat zudem die Fähigkeit, Muster aktiviert zu halten auch ohne ständige externe Anregung des Netzwerks. Das Anlegen eines Teilmusters kann nämlich bei Überschreiten einer kritischen Größe des Teilmusters zum "Feuern" (ignition) des ganzen Assemblies führen. Die aktivierten Neurone unterstützen sich dann durch die massive exzitatorische Rückkopplung zwischen den Neuronen eines Assemblies gegenseitig stark, so daß sie auch ohne Stimulus weiter aktiv bleiben (sog. Hystereseeffekt). Dies stellt ein einfaches Modell für Kurzzeitspeicherung (short-term memory, STM) in Form elektrischer Aktivität in rekurrenten Netzwerken dar (Little, 1974; Shaw et al., 1982; Amit, 1988, 1995). Langzeitspeicherung findet demgegenüber durch synaptische Plastizität in den Synapsen statt (long-term memory, LTM). In Einzelneuronableitungen hat man in verschiedenen kortikalen Arealen tatsächlich Zellen gefunden, die eine stimulusabhängige Nachaktivierung in delayed-matching-to-sample-Versuchen zeigen (Fuster und Jervey, 1980; Miyashita und Chang, 1988; Fuster, 1994). Solche Zellen könnten ein Ausdruck lokaler Kurzzeitspeicher im Sinne der Assemblytheorie sein.

Man könnte nun versuchen, den beschriebenen Assoziativspeicher als klassisches konnektionistisches Modell zu reformulieren, d.h. die gespeicherten Assemblies als Einzelknoten mit reduzierter Dynamik ("Units") aufzufassen und den Überlapp zwischen Assemblies z.B. in Form effektiver Verbindungsstärken zwischen Knoten zu modellieren. Man erhielte eine Art Graphen ähnlich einem semantischen Netz, bei dem Knoten den gespeicherten Mustern oder Assemblies entsprechen und die Gewichtung der Kanten die Ähnlichkeit zwischen den Knoten angibt. Diese Modellierungsweise ist vielleicht anschaulicher als ein Assoziativspeicher mit verteilten Repräsentationen; sie hat aber auch einige

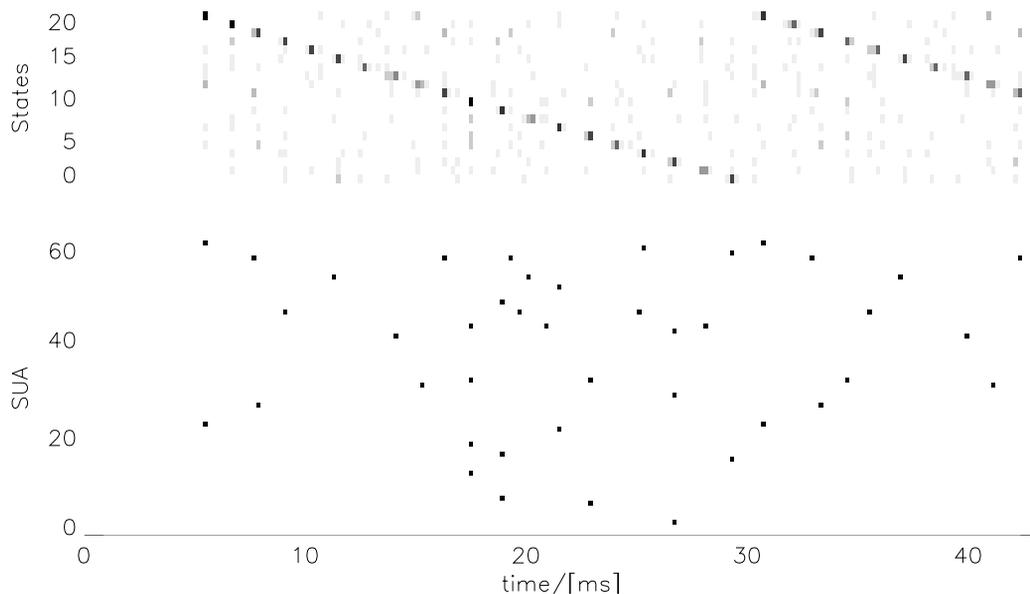


Abbildung 6: Beispiel für eine assoziativ gespeicherte zyklische Sequenz aus 20 Mustern. Im unteren Teil sind exemplarisch die Spikefolgen von 60 der insgesamt 400 simulierten Neurone gezeigt. Da die gespeicherten Muster zufällig generiert wurden, sieht man hier praktisch keine Struktur in den Spikemustern. Die geordnete, nahezu deterministische, Aktivierung der gespeicherten Muster wird jedoch sichtbar, wenn man wie im oberen Bild, den Überlapp des aktuellen Feuermusters aller Neurone mit den 20 Speichermustern (states) als Grauwertbild darstellt (weiß = kein Überlapp, schwarz = großer Überlapp). Die gespeicherte Sequenz erscheint als diagonales Band; zufällige Überlapps zwischen Speichermustern und dem Netzwerkzustand führen zudem zu weiteren Schwärzungen neben der Diagonalen.

Nachteile: man verliert beispielsweise sowohl die Fehlertoleranz und das Kategorisierungsvermögen des Assoziativspeichers, als auch seinen relativ direkten Bezug zu biologischen neuronalen Netzen. Zudem können in einem Assoziativspeicher bei dem die Muster nur wenige Einsen enthalten (sog. spärliche Kodierung) sehr viel mehr Muster gespeichert werden als er Neurone enthält (Palm, 1991). Dann wäre das entsprechende konnektionistische Netzwerk also noch nicht einmal effizienter in der Simulation als ein Assoziativspeicher mit verteilten Repräsentationen.

In der vorab besprochene Variante des Assoziativspeichers können durch sogenannte Autoassoziation statische Muster abgelegt und aus Teilen wiedergewonnen werden. Eine Erweiterung des Konzepts auf Zeitmuster ist durch sogenannte Heteroassoziation leicht möglich, indem, wie in Abbildung 2 c, jeweils Gruppen von Zellen (d.h. Muster oder Subassemblies) auf spezifische andere Gruppen konvergieren und auf diese Weise eine nahezu deterministische gerichtete Ausbreitung der Aktivität im Netz erfolgt, wenn das erste Muster in der Folge aktiviert wird. Ein Beispiel zeigt Abbildung 6.

Das simulierte Netzwerk, mit dem Abbildung 6 erzeugt wurde, bestand aus 400 Neuro-

nen. Gespeichert waren 20 Muster/Subassemblies, die je 15 aktive Neurone an zufälligen Positionen hatten. Diese 20 Subassemblies waren zyklisch in hetero-assoziativer Weise in der Kopplungsmatrix abgelegt (vgl. Abbildung 2 d). Im Bild 6 ist unten das Feuermuster von 60 der 400 Neuronen über die Zeit aufgetragen. Eine besondere Struktur ist hier nicht offensichtlich, da die gespeicherten Muster zufällig gewählt sind; rein zufälliges (z.B. Poissonsches) Feuern der Zellen würde ein ganz ähnliches Bild ergeben. Erst das obere Bild macht die deterministische Struktur des Netzwerkverhaltens deutlich. Hier ist die Ähnlichkeit des momentanen Netzwerkzustands (d.h. des momentanen Spikemusters wie im unteren Bild, nur für alle Neurone) mit den eingespeicherten Mustern (states) dargestellt. Die Ähnlichkeit ist ganz einfach gegeben durch die Anzahl übereinstimmend aktiver Neurone in einem Speichermuster und dem momentanen Netzwerkzustand. Das diagonale Band in der oberen Abbildung weist dann eine lineare Aktivierung der eingespeicherten Muster nach, die nach ca. 25 ms von vorne beginnt, da die Sequenz zyklisch geschlossen ist. Graue Bereiche neben der Diagonalen beruhen weiterhin darauf, daß die Muster zufällig erzeugt wurden und deshalb teilweise überlappen; der momentane Netzwerkzustand ist daher stets einem der eingespeicherten Muster sehr ähnlich (Diagonale), hat aber auch einen gewissen zufälligen Überlapp mit anderen Mustern.

Neuere theoretische Untersuchungen zeigen, daß sich sequentielle Feuermuster wie in Abbildung 6 in neuronalen Netzen mit zeitlich asymmetrischen Hebbischen Lernregeln durch Selbstorganisation (hier unüberwachtes Lernen) ausbilden können (Abbott und Blum, 1996; Sterratt, 1999). Solche Lernregeln verstärken eine Synapse, wenn innerhalb eines kurzen Zeitfensters (ca. 10-20ms) einem präsynaptischen Spike ein postsynaptischer folgt, und schwächen sie, wenn es umgekehrt ist. Die Existenz solcher Lernregeln in wirklichen Neuronen ist physiologisch nachgewiesen worden (Markram und Tsodyks, 1996; Magee und Johnston, 1997; Zhang et al., 1998). Möglicherweise spielen sie auch eine spezielle Rolle bei der Navigation im Raum (Blum und Abbott, 1996; Gerstner und Abbott, 1997; Samsonovich und McNaughton, 1997): im Hippokampus sind Zellen lokalisiert, die bestimmte Orte im Raum (z.B. in einem Labyrinth) repräsentieren (sog. place cells, vgl. OKeefe und Nadel, 1978). Gerichtete Übergänge der Aktivität von einem place-assembly zu einem anderen entlang Ketten determinierter Subassemblies wie in Abbildung 6 scheinen hier mögliche Pfade in der repräsentierten räumlichen Umgebung darzustellen.

Das Beispiel des Sequenzspeichers zeigt exemplarisch, daß im Assoziativspeicher nicht nur statische Muster aus Teilen wiedergewonnen werden können, sondern auch Zeitmuster. Ein bestimmter Reiz, der dem möglicherweise sogar wieder nur unvollständigen und fehlerhaften Startmuster entspricht, genügt, um eine ganze Folge von Mustern abzurufen bzw. vorherzusagen. Hierbei könnte es sich um einen Ausdruck prozeduralen Gedächtnisses handeln, um die Generierung von elementaren Bewegungen im motorischen System oder um Sprachelemente. Gerade im Hinblick auf Sprache und Bewegung sollten wir allerdings schon hier darauf hinweisen, daß das Modell wie bisher besprochen, lediglich vorgegebene lineare Sequenzen abzuspeichern in der Lage ist. Diese Struktur ist wenig flexibel und in keiner Weise ausreichend zur Erklärung des vollen Komplexitätsumfangs von Bewegungs-

, Handlungs- oder Sprachereignissen. Hier sind Erweiterungen nötig, wie wir sie später noch beschreiben werden. Für zeitliche Gedächtnisinhalte, die sich im Nachhinein ja nicht mehr ändern (sollten), mag daß Modell aber schon ganz gut passen. Weiterhin zeigen mathematische Betrachtungen, daß das Sequenzspeichermodell in dieser einfachen Form nur sehr uneffizient mit den “Hardwareressourcen” der Neuronen und Synapsen umgeht, und zwar um so uneffektiver je länger die gespeicherten Sequenzen sind (Bienenstock, 1995; Herrmann et al., 1995; Wennekens, 2000). Auch das spricht für die Notwendigkeit einer Erweiterung des Konzeptes. Wir vermuten, daß auf die angegebene Weise nur elementare Bewegungs- oder Sprachereignisse von vielleicht bis zu einigen 100ms Dauer sinnvoll gespeichert werden könnten. Darüber hinausgehend müßten diese Basisereignisse von einem (sehr wahrscheinlich) hierarchischen syntaktischen Apparat zu Äußerungsformen der für Sprache und Handlungen bekannten Komplexität zusammengesetzt werden (Cohen und Grossberg, 1986; Doya, 1999; Wickens et al., 1994).

In Abbildung 6 wirkt der Assoziativspeicher als “Generator” für ein Zeitmuster. Wir sollten erwähnen, daß man umgekehrt auch zeitliche Muster mit einem Sequenzspeicher “erkennen” kann (vgl. Abbildung 7). Hierzu erhält das Netzwerk kontinuierlich eine Eingabe, dem Strom von senso-motorischen, akustischen oder sonstigen zeitlichen Signalen entsprechend. Durch eine geeignete Schwellenkontrolle ist es dann möglich, zu erreichen, daß die Aktivierung im Netzwerk nur stabil ist, wenn die Eingabemuster und die gespeicherte Sequenz zu jedem Zeitpunkt “hinreichend” ähnlich sind. Bei zu großen Abweichungen spricht das Netzwerk nicht an. Wie ähnlich beide Zeitmuster sein müssen, kann ebenfalls durch die Schwelle kontrolliert werden, so daß man wieder eine Fehlertoleranz gegen Abweichungen erhält, nun allerdings für zeitliche Muster.

Fassen wir zusammen, was wir bisher im Rahmen des Assemblykonzeptes, bzw. seiner konkreten Realisierung in Form von Assoziativspeichernetzen, erreicht haben:

1. Die Art der Modellierung orientiert sich gleichzeitig an der Physiologie des Gehirns und an Phänomenen im Zusammenhang mit der Repräsentationen von kognitiven Ereignissen, dem Gedächtnis und dem Assoziationsvermögen.
2. Der Autoassoziativspeicher als Attraktornetzwerk kann (vor allem bei spärlicher Kodierung) eine große Zahl von Assemblies gleichzeitig enthalten, die in komplexer Weise in relationalen Beziehungen stehen dürfen. Hierbei können mittels Schwellenkontrolle bestimmte Relationen betont oder unterdrückt werden (Kontext, Aufmerksamkeit).
3. Zudem können Muster vervollständigt und klassifiziert werden (Assoziativität und Kategorienbildung, Stimuluskonstanz) oder Subassemblies zu größeren Strukturen zusammengefaßt werden (Kompositionalität). Das “Bindungsproblem” kann hierbei einerseits durch verstärkte synaptische Konnektivitäten, andererseits durch synchronisierte neuronale Aktivität gelöst werden.

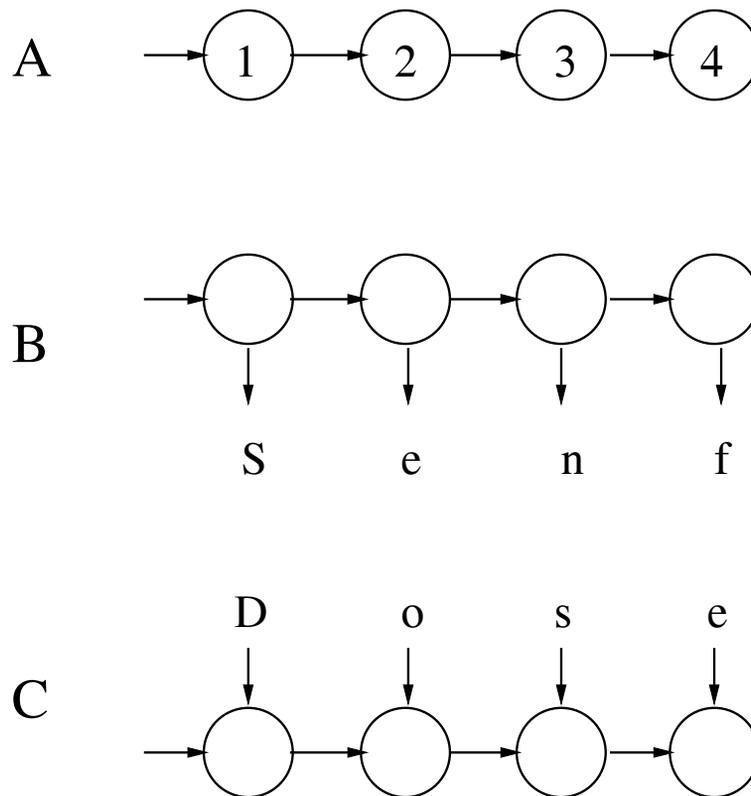


Abbildung 7: A) Schema des Sequenzspeichers. Die Speichermuster sind als symbolische Knoten in linearer Ordnung dargestellt. B) Wenn jeder Knoten zusätzlich noch (assoziative) Ausgangsverbindungen hat, erhält man einfache Mustergeneratoren. C) Analog führen zusätzliche assoziative Eingangsverbindungen zu Mustererkennern. Die gespeicherte Sequenz wird bei passender Schwelle nur vervollständigt, wenn zu jedem Zeitpunkt das Feuern der Vorgängermusters erfolgt und das korrekte Eingangsmuster anliegt.

4. Neben statischen Entitäten lassen sich zeitliche Eigenschaften und Relationen im Sequenzspeicher repräsentieren. Dies ermöglicht das Generieren und Erkennen einfacher zeitlicher Handlungen und Signalfolgen. Zudem sind “Prädiktionen” von bestimmten Stimuluskonstellationen auf Folgezustände möglich (Reiz-Reaktions-Schemata).
5. Hysterese stellt einen einfachen Mechanismus für Kurzzeitspeicherung in den neuronalen Erregungsmustern bereit. Durch die wechselseitige starke Verknüpfung zwischen den Neuronen eines Assemblies kann sich die Aktivität im Netzwerk ohne fortlaufende Aktivierung durch externe Reize aufrecht erhalten. Neuronale Aktivität “fließt” entlang verstärkter Verbindungsbahnen in komplexen neuronalen Repräsentationen mit statischen und zeitlichen Komponenten.
6. Über Hebb’sche synaptische Plastizität, lokal an den einzelnen Synapsen, induziert die neuronale Aktivität zudem eine Form des Langzeitgedächtnisses. Kurz- und Langzeitgedächtnis stellen sich hiernach als verschiedene physiologische Eigenschaften in ein und demselben Netzwerk dar; Kurzzeitspeicherung als elektrisches neuronales Erregungsmuster, Langzeitspeicherung als Muster der synaptischen Konnektivitäten.

4 Globale Netzwerke und Interaktion von Assemblies

Bisher bezogen sich unsere Ausführungen nur auf “lokale” Netzwerke. Hiermit ist gemeint, daß wir wenig über die anatomische Struktur des Gehirns im Großen gesagt haben und damit über die Interaktionsmöglichkeiten zwischen lokalen (Sub-)Assemblies über weite Distanzen im Gehirn. Selbstverständlich setzt die Anatomie Randbedingungen für die Modellierung und das Verständnis der Informationsverarbeitungsprozesse im Gehirn. Wichtige Aspekte hierbei sind:

1. Die Aufteilung des Cortex in eine Vielzahl separater Areale mit jeweils spezifischen Aufgaben, sowie die starke, oft bidirektionale, wechselseitige Vernetzung cortikaler und subcortikaler Bereiche, die mit ähnlichen Aufgaben befaßt sind.
2. Die Existenz einer Reihe funktioneller Systeme im Gehirn, die bestimmte spezielle kognitive Aufgaben unterstützen.
3. Das Zusammenwirken cortikaler Areale bei der Erfüllung kognitiver Aufgaben, die über simple Reiz-Reaktions-Schemata hinausgehen.

Mit diesen drei Problemkreisen wollen wir uns im folgenden Teil unserer Darstellungen auseinandersetzen.

Betrachtet man die Anatomie und Physiologie des Gehirns im Großen, ist ein wichtiges Bauprinzip die Existenz mehrerer sensorischer und motorischer “Hierarchien” im Neocortex (Lashley, 1950; Fuster, 1994; Young et al., 1995), etwa die visuellen, auditorischen und sensomotorischen Areale auf der Eingangsseite, oder motorische und Sprachareale auf der Ausgangsseite. Diese afferenten und efferenten Kanäle arbeiten parallel und sind selbst wieder in verschiedene Areale unterteilt, die sich mit jeweils spezifischen Teilaufgaben befassen (z.B. Livingstone und Hubel, 1987, 1988). Areale, die verwandte Aufgaben bearbeiten, sind hierbei intensiv und oft in beide Richtungen über synaptische Projektionen miteinander verbunden - insofern ist der Begriff “Hierarchie” vielleicht etwas irreführend (Swindale, 1990; Bullier et al., 1993; Felleman und van Essen, 1991; Merigan und Maunsell, 1993). Er beruht neben anatomischen Gesichtspunkten in erster Linie darauf, daß das Antwortverhalten von kortikalen Neuronen zunehmend komplexer wird, je weiter entfernt (etwa gemessen in der Anzahl synaptischer Umschaltstufen) man sich von den primären sensorischen oder motorischen Arealen befindet. Beispielsweise sind visuelle Neurone im primären Sehcortex (V1) sensitiv für bewegte orientierte Kanten, Farbe, Kontrast, usw. (sog. rezeptive Felder, Hubel und Wiesel, 1962), in höheren Arealen findet man dagegen Zellen, die spezifisch auf recht komplexe figurale Bildelemente reagieren, auf Gegenstände und sogar auf Gesichter (Miyashita und Chang, 1988; Tanaka, 1993, 1997; Logothetis, 1996; Martin et al., 1996; Damasio et al., 1996). Dasselbe gilt ganz ähnlich auch für andere sensorische Pfade, etwa das auditorische System (Rauschecker, 1998) oder das somato-sensorische (Fuster, 1994; Iwamura, 1998). Sogar im motorischen System ändert sich die Komplexität der Zelleigenschaften vom Konkreten zum Abstrakten hin, wenn man von primären Arealen zu “höheren” geht: Primäre motorische Areale sind somatotopisch organisiert, ihre Neurone kodieren konkrete Motoraktionen; in sekundären motorischen Arealen kombinieren die Zellen z.B. Informationen verschiedener Sinnesmodalitäten zur Ausführung einer Aktion, oder sie sind an der zeitlichen Ordnung einer Sequenz von Aktionen beteiligt (Fuster, 1994).

Im Rahmen der konnektionistischen Modellierung der Gehirnfunktion haben diese biologischen Befunde zum Bild des Cortex (besser gesagt seiner sensorischen Komponenten) als eines hierarchisch organisierten Musteranalysesystems geführt: Primäre Areale detektieren einfache Merkmale in Sinnesdaten, geben sie an die nächste Stufe weiter, wo sie zu komplizierteren Merkmalen synthetisiert werden, usw. Entsprechende Modelle simulieren hierzu sensorische Areale als mehrstufige Karten für zunehmend kompliziertere Merkmale z.T. bis schließlich hin zu Buchstaben oder Objekten (z.B. Tononi, 1992; Stöcker et al., 1996; Fukushima und Myaki, 1982; Rao und Ballard, 1997).

Aus der Sicht der Assemblytheorie ergeben sich jedoch weitere Fragen.

1. Welchen Funktion haben die massiven bijektiven Projektionen zwischen Arealen und die damit verbundene Konvergenz, Divergenz und Rekurrenz?
2. Werden Signale in sensorischen Filterhierarchien einfach nur umkodiert bis auf höchster Ebene schließlich eine komprimierte Darstellung erreicht ist, auf der allein

(bewußte) kognitive Prozesse operieren? Oder wird umgekehrt Information jeweils auch auf allen Ebenen lokal repräsentiert, bearbeitet und gespeichert?

3. Komplexe Entitäten werden im Gehirn über eine Vielzahl von Arealen mit jeweils spezifischen Teilaufgaben verteilt repräsentiert. Wie wird diese weit verteilte Information konzertiert verarbeitet (Bindungsproblem, vgl. Treisman, 1996)?

Zu diesen Fragen gibt es eine Reihe älterer und neuerer Arbeiten, die sich mit der detaillierten Interaktion zwischen zwei oder mehreren cortikalen Arealen befassen (Damasio, 1989; Tononi et al., 1992; McGregor, 1993; Bibbig et al., 1995, 1996a; Renart et al., 1999ab; Sommer und Wennekers, 1998, 2000; und mehr). Gerade die neueren Arbeiten sind hierbei oft vom Konzept des “Bindens durch Synchronisation” beeinflusst.

Sicherlich ist das Konzept der Filterhierarchie eines, daß einen wichtigen Aspekt der Informationsverarbeitung gut erfaßt, nämlich die verschiedenen Abstraktionsebenen neuronaler Antworteigenschaften, wie sie auch physiologisch nachweisbar sind. Diese dürften tatsächlich durch divergent-konvergente Vorwärtsverbindungen zu großen Teilen erklärbar sein. Allerdings werden die massiven lateralen und Rückkoppelverbindungen im Cortex praktisch vernachlässigt. Zudem verfügt der Assoziativspeicher noch über die Kurzzeitspeichereigenschaft in Form rekurrenter Aktivität. In der Assemblytheorie kann man alle diese Eigenschaften integrieren, wenn man berücksichtigt, daß Muster nicht wie im oben beschriebenen technischen Modell völlig zufällig sind, sondern besonders in den sensorischen primären Arealen natürlich sehr stark von den Sinnesdaten bestimmt werden. Es sind eine Reihe von Modellen dafür entwickelt worden, wie sich in Gegenwart synaptischer Plastizität, durch Korrelationen, d.h. immer wiederkehrende Strukturelemente in den Sinnesdaten, entsprechende rezeptive Felder in sensorischen Arealen ausbilden können, die diese Elemente detektieren (Review in Erwin et al., 1995). Sind primäre Felder ausgebildet, kann sich der Prozeß in weiteren, höheren Arealen fortsetzen. Dort entstehen dann Zellen mit komplexeren Eigenschaften, usw. Dies entspricht der Ausbildung einer Merkmalshierarchie von peripheren zu zentralen Arealen durch Selbstorganisation. Allerdings werden im Gehirn selbst bei Reizung mit einem einfachen Lichtbalken in einem Areal stets relativ viele Zellen aktiviert; es kommt also letztendlich nicht zur Bildung von Einzelneuronen, die repräsentativ für ganze Objekte stehen (sog. Großmutterneurone), sondern auch in den höchsten Verarbeitungsstufen werden Stimuli stets durch die Aktivierung ganzer Pools von Zellen charakterisiert (Populationskode). Diese Pools können als Assemblies aufgefaßt werden, wobei man allerdings die Annahme machen muß, daß die entsprechenden Zellen untereinander stark verknüpft sind. Das ist jedoch plausibel, da die gleichzeitige Aktivierung der Zellen auch lokal im Areal die synaptischen Verbindungen stärken sollte. Auf diese Weise können lokale Assemblies in höheren Arealen entstehen, die nun auch nicht mehr so sehr durch die Topographie der Eingabedaten bestimmt sein müssen, sondern physikalisch mehr oder weniger zufällig über das lokale Netzwerk verteilt sein können. Letzteres als Folge der hohen Divergenz cortiko-cortikaler Projektionen. Dies

paßt zu dem biologischen Befund, daß die retinotopische Ordnung im visuellen oder auch die tonotopische Ordnung im auditorischen System in höheren Arealen verloren geht.

Auf die vorangehend dargestellte Weise können die Eigenschaften von Merkmalshierarchien und Assoziativspeichern in der Assemblytheorie integriert werden. Interessanterweise scheint es zu sein, daß Zellen mit typischer Kurzzeitspeichereigenschaft (vgl. Hystereseeffekt im Assoziativspeicher) im visuellen System erst auf relativ hoher Ebene auftreten; vorgeschaltete Stufen passen dort besser zum Filterkonzept. In senso-motorischen und motorischen Arealen ist das aber nicht mehr der Fall; dort sind Speicherphänomene auch schon in den primären Arealen vorhanden (Miyashita und Chang, 1988; Fuster, 1994). Dies legt nahe, daß auf nahezu allen Ebenen Information nicht nur analysiert oder umkodiert wird, sondern daß das jeweilige Areal auch direkt an der Repräsentation und Speicherung beteiligt ist. Fein aufgelöste und konkrete Details sollten dann in sensornahen Arealen bearbeitet werden, komplexere Merkmale und Zusammenhänge, abstrakte Klassen, Kategorien, Kontexte, etc. in höheren Arealen. Dies bestätigt sich in einer Vielzahl von Läsionsstudien, wonach lokalisierte cortikale Schädigungen oft hochgradig spezifische kognitive Fehlleistungen und Ausfälle zur Folge haben. Es gibt demnach im Gehirn anscheinend kein besonderes Areal für "höhere Hirnleistungen". Spezielle Leistungen ergeben sich jeweils aus den vielfältigen Interaktionen zwischen den beteiligten Arealen mit ihren spezifischen Teilaufgaben. Die häufig anzutreffenden bijektiven Projektionen zwischen Hirnbereichen sollten diese komplexen Interaktionen unterstützen.

Viele Teile des hier gezeichneten Bildes wurden bereits in Computermodellen untersucht. Im einfachsten Fall werden dabei nur zwei Areale betrachtet. Einige typische Beispiele sind im folgenden aufgeführt.

1. In sensorischen Systemen kann der Einfluß von Rückwärtsverbindungen in topographisch gekoppelten Schichtenmodellen mit Kartenstruktur untersucht werden. Hier führt die Aktivierung von Merkmalsneuronen in einem höheren Areal durch die Rückprojektionen zu einem stärkeren Ansprechen im vorgeschalteten Areal. Dies kann als "Binden" der entsprechenden Zellgruppen durch erhöhte Aktivität aufgefaßt werden (Fukushima und Myaki, 1982; Neumann und Mössner, 1996; Grossberg, 1999ab).
2. Ganz ähnlich sind "top down"-Aufmerksamkeitsmodelle aufgebaut, nur daß die Aktivierung in den niedrigeren Arealen hier nicht unbedingt eine Voraussetzung für die im höheren Areal ist. Das höhere Areal moduliert die Aktivität relevanter Bereiche/Zellen im vorgeschalteten durch Schwellenkontrolle. Modelle für Aufmerksamkeit findet man in (Posner, 1990; Desimone und Duncan, 1995; Kirkland und Gerstein, 1999; Review in Suder et al., 2000).
3. Die "Adaptive Resonanztheorie" (ART) von Carpenter und Grossberg läßt sich ebenfalls im Sinne von 1. (und 2.) auffassen (Carpenter und Grossberg, 1991;

Grossberg, 1999ab). Hier kommt aber noch hinzu, daß durch einen Selbstorganisationsprozeß auf der höheren Ebene bei “neuen” Eingangsmustern auch neue “Units” zur Repräsentation gebildet werden.

4. Auf der Beschreibungsebene der Assoziativspeicher ist wiederholt bidirektionales Retrieval zwischen zwei und mehreren Arealen betrachtet worden (Bibbig et al., 1995, 1996a; Renart et al., 1999ab; Sommer und Wennekers, 1998, 2000; Sommer und Palm, 1999). Hier unterstützt das Anlegen (externer Input) oder Ansprechen (Feuern) eines Musters die Erkennung des zugehörigen Musters im anderen Areal. Kognitiv relevante Phänomene in diesem Zusammenhang sind z.B. crossmodales Retrieval oder auch die Absenkung von Wahrnehmungschellen bei gepaarten Reizen unterschiedlicher Modalität.

Wenn das Zusammenspiel von mehr als zwei Arealen miteinander untersucht wird, werden die möglichen qualitativen Verhaltensweisen schnell unübersichtlich. Derartige Modelle orientieren sich daher in der Regel an ganz konkreten kognitiven Leistungen und verwenden Modelle mit einer speziellen Struktur, die die entsprechenden Subsysteme des Gehirns darstellt. Hierzu gehören (u.a.) Aufmerksamkeitsmechanismen (Kirkland und Gerstein, 1999; Suder et al., 2000), die Wechselbeziehung zwischen Cortex und der Hippokampusformation im Hinblick auf Navigation im Raum (Brown and Sharp, 1995; Burgess et al., 1997; Franz und Mallot, 2000; Lieblich und Arbib, 1982; Qin et al., 1997), Gedächtniskonsolidierung und den Einfluß von Emotionen auf cortikale Informationsverarbeitung (Miller, 1991; Bibbig et al., 1995, 1996ab; McClelland et al., 1999), die Wechselwirkung von Neocortex, Kleinhirn und Basalganglien bei der Bewegungskontrolle (Gross et al., 1999), und ähnliches.

Aus dem Gesagten ergibt sich, daß Assemblies in der Regel nicht in einem einzigen Areal oder sogar nur einer einzigen Kolumne lokalisiert sein werden, selbst dann nicht, wenn es sich bei dem repräsentierten Ding um eine relativ einfache Sache handelt. Wie am visuellen System deutlich wird, zerlegt das Gehirn auch schon einfache Objekte in ihre Eigenschaften, die mehr oder weniger getrennt in verschiedenen Cortexarealen verarbeitet werden (Form, Farbe, Bewegung, etc., vgl. Livingstone und Hubel, 1987, 1988; Felleman und van Essen, 1991). Die Einzelaspekte mögen dabei in den speziellen Arealen durch Gruppen stark verbundener Zellen dargestellt sein, das ganze Objekt wird jedoch verteilt repräsentiert. Man kann dann von “globalen Assemblies” sprechen. Modelle und Ansätze, die aus diesem Blickwinkel interpretiert werden können, findet man beispielsweise bei Hebb (1949), Wickelgren (1969, 1981, 1992), Milner (1974, 1996), Braitenberg (1978), Palm (1982, 1990), Amit (1988, 1995), Damasio (1989), Gerstein et al. (1989), Pulvermüller (1992, 2000), McGregor (1993), Bienenstock (1995), Wennekers und Palm (2000), u.a.

An dieser Stelle sollten wir noch einmal auf ein Problem zu sprechen kommen, das wir schon früher erwähnt haben, und daß sich besonders in derart verteilten Repräsentationen stellt, wie es “globale Assemblies” sind. Dies ist das Bindungsproblem oder die

Frage, ob und wie zusammengehörige Teile im Gehirn auch als solche gekennzeichnet werden. So etwas scheint intuitiv notwendig, wenn man sich klar macht, daß im Gehirn zwar meistens wohl nur jeweils ein Sinnzusammenhang repräsentiert ist, aber dabei oft mehr als nur ein Gegenstand gleichzeitig präsent ist und man trotzdem eine einheitliche und unvermischte Wahrnehmung verschiedener diskreter Dinge hat. Dies ist etwa der Fall, wenn man sich in einer Situation in einem Raum mit mehreren gerade wichtigen Personen und Gegenständen befindet, oder wenn man sich dabei auch noch eine Abfolge von Argumenten oder Handlungen zurechtlegt oder eine Telefonnummer memoriert. Genauere Ausführungen zur möglichen Komplexität der Inhalte unseres Arbeitsgedächtnisses finden sich bei Baddeley (1996).

In einfachen konnektionistischen Ansätzen wird jedem modellierten Gegenstand genau eine "Unit" zugeordnet. Hier tritt das Bindungsproblem scheinbar gar nicht auf. Tatsächlich stoßen solche Ansätze aber schnell an Grenzen, nämlich dann, wenn eine Vielzahl von Entitäten repräsentiert werden muß, und diese - wie in der wirklichen Welt - in ganz unterschiedlichen Erscheinungsformen vorkommen können. Für jede Form wäre eine besondere Unit notwendig, sonst könnte man einen schlanken Mahagonispazierstock und einen Knüppel mit Griff nicht auseinander halten; der Unterschied wäre nicht repräsentierbar. Dies würde zu einer ungeheuren Zahl an "Units" führen, die für die Darstellung der Welt notwendig wären (kombinatorische Explosion).

Das Problem hängt natürlich mit der Kompositionalität von Dingen zusammen, mit ihrer (noch nicht einmal eindeutigen) kognitiven Zerlegbarkeit in Eigenschaften. Diese drückt sich wiederum im Gehirn in der Art der Informationsverarbeitung in spezifischen Pfaden und Arealen aus. Gerade diesen Aspekt erfaßt das Assemblykonzept in direkter Weise. Das Bindungsproblem wird dann oft in der folgenden Frage formuliert: Angenommen wir sehen gleichzeitig auf ein rotes Dreieck und ein grünes Quadrat. Nun werden aber doch die Farben (idealisierter Weise) unabhängig von den Formen in verschiedenen Arealen repräsentiert und verarbeitet. Wie kommt es, daß das Gehirn die Zellen für "Röte" im Farbareal denen für "Dreieck" im Formareal zuordnet und nicht denen für "Quadrat". Dies zumal es ja auch grüne Quadrate gibt und sogar welche in allen Farben. Durch verstärkte Verbindungen kann das nicht gehen, denn dann müßten die Verbindungen von allen Farben zu allen Formen stark sein und man dreht sich im Kreise. (Derartige verstärkte Verbindungen nimmt das Assemblykonzept allerdings für Repräsentationen eines Dings an, wenn es typischerweise immer wieder mit bestimmten Eigenschaften auftritt; dann können die entsprechenden Verbindungen durch synaptische Plastizität verstärkt werden.)

Ein dynamischer Mechanismus scheint notwendig und einen solchen glaubt man im Konzept des "Bindens durch Synchronisation" gefunden zu haben, gemäß dem Neurone, die zusammengehören, synchronisierte Aktivität aufweisen und asynchron zu Neuronen anderer Repräsentationen sind (von der Malsburg, 1981). Speziell wird vermutet, daß cortikale Gammaoszillationen das Binden unterstützen. Hierfür gibt es Evidenz aus Mikroelektrodenableitungen in verschiedenen cortikalen Arealen von Katzen und Affen

(Eckhorn et al., 1988; Eckhorn, 1999; Singer und Gray, 1995) und neuerdings auch aus EEG-Ableitungen bei Menschen, wo sogar ein Zusammenhang mit der bewußten Wahrnehmung und dem assoziativen Lernen auf psychologischer Ebene zu bestehen scheint (Rodriguez et al., 1999; Miltner et al., 1999).

Seit der Entdeckung stimulusabhängiger Synchronisation im visuellen Cortex, inspiriert durch den Zusammenhang zum Bindungsproblem und einige interessante physikalisch-mathematische Probleme, ist eine immense Zahl theoretischer und numerischer Modellstudien durchgeführt wurden, die sich mit der Synchronisation neuronaler Aktivität auseinandersetzen (z.B. Erb und Aertsen, 1992; Tononi et al. 1992; Gerstner und van Hemmen, 1993; Gerstner et al., 1993; Schillen und König, 1994; Bibbig et al., 1995, 1996a; Stöcker et al., 1996; Jensen und Lisman, 1996; und viele mehr; Übersichtsartikel in Aertsen und Arndt, 1993; Gerstner, 1995; Wennekers und Palm, 1997, 2000). Hierbei zeigt sich, daß schon relativ schwache laterale cortexintrinsiche Verbindungen oder bijektive Projektionen zwischen Arealen neuronale Synchronisation unterstützen. Unklar ist allerdings immer noch die genaue Ursache von Gammaoszillationen und ob Synchronisationen eher von exzitatorischen oder inhibitorischen Verbindungen vermittelt wird. Ebenfalls ist noch die Frage offen, ob verschiedene gleichzeitig aktive Objektrepräsentationen völlig unkorreliert sind oder sich durch synchrones Feuern der jeweiligen Assemblies zu gewissen Zeiten innerhalb der kollektiven Gammaoszillation auszeichnen (Phasenkodierung). Verschiedene Assemblies würden hiernach in zeitlich festgelegter Reihenfolge feuern.

Synchronisierte Aktivität ist ein Mechanismus um lokale Assemblies, die untereinander nur schwach verbunden sind, weil sie vielleicht nur gelegentlich gemeinsam auftreten, zusammenzubinden (vgl. Formen, die in jeder Farbe auftreten können). Andere Mechanismen erscheinen aber möglich, die das Problem ebenfalls lösen. Hierzu gehört beispielsweise transiente synaptische Plastizität auf einer relativ schnellen Zeitskala, die nur vorübergehend die entsprechenden Zellgruppen miteinander koordiniert. Zudem kann man argumentieren, daß das Bindungsproblem im Gehirn gar nicht in der krassen Form auftritt, wie es das Beispiel oben suggeriert. Dies in erster Linie deshalb, weil es unplausibel ist, daß es beispielsweise genau ein fest umrissenes Assembly für "Röte" gibt. Tatsächlich gibt es farbsensitive Zellen in vielen Arealen und diese Zellen reagieren in aller Regel auch noch auf andere Stimuluseigenschaften, wie die Bewegung oder den Ort eines Objekts. Daß heißt, die sensorischen Pfade im Gehirn sind gar nicht so strikt separiert, wie das obige Beispiel nahegelegt. Das wiederum bedeutet, in praktisch jedem Assembly sollten Zellgruppen zu finden sein, die es ermöglichen alleine aufgrund ihrer Aktivierung zwischen zwei oder mehreren Stimuli zu unterscheiden. Solche Zellen wird man zweifelsohne beispielsweise im primären Sehcortex finden, weil dort gleichzeitig dargebotene Stimuli zumindest teilweise räumlich getrennt sind und mit einer relativ feinen räumlichen Auflösung verarbeitet werden. Ausgehend von solchen Zellen, die Eigenschaften (hier Ort und Farbe) von Objektmerkmalen kombinieren sind Mechanismen denkbar, die auch aufgrund reiner Ratenmechanismen ohne Synchronisation das Bindungsproblem lösen können (Phillips und Singer, 1997; Palm und Wennekers, 1997). Nur weitere Experimente können klären,

wie neuronales Binden wirklich vonstatten geht.

Wir kommen nun abschließend zu der Frage, wie cortikale Assemblies wenigstens im Prinzip bei der Erfüllung kognitiver Aufgaben zusammenwirken können, die über einfache Assoziationen oder Reiz-Reaktions-Schemata hinausgehen. Hierbei spielen Pläne und Gedächtnis besonders von zeitlichen Beziehungen zwischen Stimuli, internen Ereignissen und Verhaltensreaktionen eine Rolle, es geht also um die Integration von Informationen und Prozessen in allen Teilen der sensorischen und motorischen Peripherie eines Individuums und um die Koordinierung komplexer Verhaltensmuster.

Bisher haben wir uns mit den Möglichkeiten der Repräsentation räumlicher und einfacher zeitlicher Eigenschaften von Entitäten in Assemblies beschäftigt, die über die verschiedenen sensorischen und motorischen Hierarchien verteilt sind. Hier war ein wichtiger Aspekt, daß die Arten der Repräsentation vom Konkreten zum Abstrakten fortschreitend in neuronalen Karten angelegt sind, aufsteigend in sensorischen Arealen, absteigend in motorischen. Wie aber wird die Sensorik mit der Motorik verbunden (senso-motorische Schleife)?

Zunächst ist zu sagen, daß bereits auf niederen Ebenen zahlreiche wechselseitige Verbindungen zwischen sensorischen und motorischen Bereichen im Gehirn existieren. Das reicht von simplen Reflexbögen im Rückenmark, über die direkte Kopplung zwischen dem primären sensomotorischen und motorischen Areal (die benachbart liegen), bis hin zu Zellen, etwa in parietalen Arealen, die visuelle und akustische Information zur Ausführung bestimmter Aktionen in Verhaltensexperimenten integrieren. Es gibt demnach nicht nur eine "senso-motorische Schleife", sondern viele auf allen Abstraktionsebenen der Verhaltenskontrolle. Die genauen Funktionen dieser Schleifen sind allerdings bisher nur sehr unzureichend bekannt. Modellierungsansätze gibt es daher nur relativ wenige, von denen viele zudem nur Teilaspekte betrachten und weniger das Zusammenwirken ausgedehnter cortikaler Systeme bei der Motorkontrolle (z.B. Georgopolous, 1986; Milner, 1992; Wickens et al., 1994; Pasemann, 1995; Salinas/Abbott, 1995; Burdet und Milner, 1998; Gross et al., 1999).

Neben diesen Wechselwirkungen auf niederen Ebenen ist weiterhin bekannt, daß die sensorischen und motorischen Pfade anatomisch mit den Arealen des sogenannten frontolimbischen Komplexes verbunden sind (Fuster, 1994; Young et al., 1995). Hierbei handelt es sich um bevorzugt im frontalen Bereich des Neocortex und der Hippokampusformation gelegene Strukturen, die Information aus weiten Bereichen des Gehirns integrieren und anscheinend mit der Koordinierung von Lernen und Verhalten auf einem hohen Abstraktionsniveau betraut sind. Einige dieser Bereiche sind für spezielle Aufgaben wie Navigation im Raum, Emotionen und Gedächtniskonsolidierung zuständig. Konnektionistische Modelle hierzu haben wir oben schon erwähnt. Ein spezielles Areal in diesem Bereich ist der präfrontale Cortex. Verhaltensexperimente legen nahe, daß Neuronen in diesem Areal in besonderer Weise mit dem Gedächtnis räumlicher und zeitlicher Beziehungen von verhaltensrelevanten Ereignissen befaßt sind (Fuster, 1973, 1989, 1994). Dies wird beispielsweise in Experimenten deutlich, in denen ein Versuchstier (typischerweise ein Affe)

einen Stimulus angeboten bekommt und nach einer gewissen Zeit ohne Stimulus entscheiden muß, ob ein dann angebotener Testreiz dem ersten Reiz entspricht (delayed-match-to-sample task). Die Gleichheit muß durch eine Verhaltensreaktion signalisiert werden, die Ungleichheit durch eine andere Reaktion (z.B. das Drücken von Knöpfen oder Hebeln). Ableitungen von Neuronen im präfrontalen Cortex während solcher und ähnlicher Aufgaben können ganz spezifisches Antwortverhalten entsprechend der jeweiligen Versuchsbedingung aufweisen (match vs. non-match etc.). Zum einen findet man Zellen mit Kurzzeitspeichereigenschaft, die nur während der Delayperiode aktiv sind. Diese Zellen können zudem selektiv für den angebotenen Stimulus sein (sog. cue-cells, vgl. Fuster, 1994). Außerdem kann sich ihr Feuerverhalten drastisch ändern, je nachdem, ob der jeweilige Stimulus verhaltensrelevant ist oder nicht (um dies zu testen muß die Aufgabe allerdings etwas komplizierter sein als das delayed-matching-to-sample Paradigma; gewisse Stimuli müssen eben manchmal, aber nicht immer, relevant sein). Auf der anderen Seite findet man auch Neurone, die selektiv die auszuführende Motorantwort kodieren (motor set cells). Sie zeichnen sich durch eine während der Delayphase zunehmende Feuerrate aus, quasi in Vorwegnahme der Motoraktion. Die Aktivität ist zudem abhängig von der Vorhersagbarkeit der erforderlichen Antwort; sie ist größer, wenn auf einen bestimmten Stimulus mit hoher Wahrscheinlichkeit stets dieselbe Response auszuführen ist. Entsprechend verkürzen sich Reaktionszeiten. Die Motoraktion selbst wird aber erst bei Präsentation des Teststimulus ausgeführt.

Diese Befunde legen nahe, daß es im Präfrontalcortex ein Art von Repräsentationen für Handlungsschemata gibt. Bestimmte verhaltensrelevante Ereignisse können dort selektiv Aktivität in den Zellen hervorrufen. Diese können die Information über einen Stimulus einige Zeit speichern, eine adäquate Antwort vorhersagen und sie in Abhängigkeit vom Teststimulus auslösen. Wie kann man sich dies aus der Sicht der Assemblytheorie genauer vorstellen?

Zunächst ist zu betonen, daß die Ausbildung von Zellen im Präfrontalcortex, die an der Koordination von komplexen Aufgaben beteiligt sind, tendenziell verhaltensspezifisch erfolgt. Zellen, die einfach nur die Präsenz eines Reizes in der äußeren Welt oder eine Reaktion darauf repräsentieren und speichern können - unabhängig von der momentanen Relevanz - gibt es ja auch bereits in höheren Arealen der sensorischen und motorischen Bahnen, exemplarisch etwa im Inferotemporalcortex etc. (Fuster und Jervey, 1980; Miyashita und Chang, 1988; Logothetis, 1996). Diese Areale sind zudem mit frontalen Arealen wechselseitig verknüpft. Vereinfachend kann man daher sagen, daß die Verhaltensspezifität in den Antwortmustern präfrontaler Zellen nahelegt, daß bestimmte im jeweiligen Kontext relevante Stimulationsbedingungen in globalen Assemblies verteilt über sensorische Bereiche repräsentiert werden, daß sich aber zusätzlich (wahrscheinlich) im Präfrontalcortex ein lokales Assembly bildet, dessen Aktivierungszustand gewissermaßen nichts weiter als das Auftreten des entsprechenden verhaltensrelevanten Ereignisses signalisiert; die genauen Details des Stimulus werden an anderer Stelle abgelegt. Entsprechend würde diese präfrontale Gruppe von Zellen zwar Informationen aus weiten Bereichen des Cortex

integrieren (auch multimodal), diese aber nicht selbst wieder beinhalten. Dieses Assembly hätte sich außerdem erst während des Trainings des Affen wohl durch Hebb'sches Lernen gebildet. Es ist danach ein abstrakter Schlüssel, der ein Verhaltensereignis signalisiert, und über den (wenn notwendig) über Rückprojektionen in andere Areale auf Details der entsprechenden Stimulussituation geschlossen werden kann oder aber eine Handlung ausgelöst werden kann. Computersimulationen, die zeigen, wie man sich die Ausbildung solcher komprimierter multimodaler Cue-Assemblies oder Chunks durch synaptische Plastizität vorstellen kann, findet man in Bibbig et al. (1995, 1996a; vgl. auch Miller, 1991; Wickelgren, 1992; Palm, 1993ab). Aspekte der Rückadressierung von Stimulusdetails aus solchen Schlüsseln werden zudem durch bi- oder multimodale Assoziativspeicher gut erfaßt (Bibbig et al., 1995, 1996a; Sommer und Wennekers, 1998, 2000; Renart et al., 1999ab). Einer dieser Speicher würde die präfrontalen Subassemblies enthalten, die anderen, in verteilter Weise, das globale Assembly, das den Stimulus (und evtl. seine kontextuelle Einbettung) genau charakterisiert.

Ganz analog zu derartigen Cue-Assemblies sollten auch Motor-Set-Assemblies existieren, die auf hoher Ebene, entsprechend abstrakt und detailarm, für die Verfügbarkeit gewisser Reaktionsmuster stehen. Dementsprechend käme dem Präfrontalcortex die Aufgabe zu, zwischen korrespondierenden Cue- und Motor-Set-Assemblies in einer komplexen kognitiven Aufgabe zu vermitteln, also Kombinationen von Reizsituationen und darauf erwartete Antworten assoziativ zu bilden und bei Vorliegen eines Stimulus die entsprechende Reaktion einzuleiten. Die genauen Verhaltensmuster und Bewegungen wären auch hier nicht selbst präfrontal abgelegt, sondern in Motorarealen und diese unterstützenden Bereichen. Zudem würde die konkrete Ausprägung einer Verhaltensreaktion durch die senso-motorischen Schleifen auf niederen Ebenen kontrolliert.

Zwei Eigenschaften von cue- und motor-set-Zellen, wie sie experimentell beobachtet werden, sind hierbei noch nicht berücksichtigt: Dies ist zum einen die zeitlich anwachsende Voraktivierung von motor-set-Zellen während der Delayphase und die Abhängigkeit der Aktivität von der Vorhersagbarkeit der Response. Zweitens lösen motor-set-Zellen erst eine Reaktion aus, wenn sie nach der Delayphase vom Teststimulus quasi getriggert werden. Die Beziehungen sind also etwas komplizierter, als im einfachen Assoziativspeicher. Computersimulationen, welche die unterliegenden Prozesse nachbilden, sind uns noch nicht bekannt; ein Modellansatz könnte allerdings wieder auf dem Assoziativspeicherkonzept aus spikenden Neuronen aufbauen, wo das zeitliche Anwachsen der Feuerraten durch eine Schwellenmodulation erreicht werden kann, entsprechend vielleicht einem Aufmerksamkeitsprozeß oder einer Art "inneren Uhr", welche die verstrichene Zeit während der Delayphase mißt. Die Abhängigkeit der Aktivität von der Vorhersagbarkeit der Reaktion sollte sich zudem aus einfachen Hebb'schen Lernregeln ergeben, wenn ein Stimulus-Response-Paar häufiger gemeinsam auftritt als ein anderes. Entsprechend verstärkte Verbindungen zwischen den Cue- und Motor-Set-Assemblies würden proportional die Feuerraten im Motor-Set-Assembly regulieren. Die Auslösung der Motorantwort erfordert weiterhin eine Kombination der präfrontalen Aktivität während der Delaypha-

se mit dem vom Teststimulus aktivierten globalen Assembly. Dies kann man wieder als Assoziation auffassen, wobei die präfrontale Aktivität alleine noch nicht ausreicht, um die Verhaltensreaktion auszulösen, sondern zum Erreichen der Aktivierungsschwellen in den höheren Motorarealen noch zusätzlich die vom Teststimulus hervorgerufene Aktivität benötigt wird. Plausibel ist hierbei auch die Annahme, daß allein die plötzliche Präsenz des Teststimulus noch einen weiteren Schwellensenkungsprozeß einleitet, der letztendlich erst das Verhalten auslöst. Je nachdem, ob der primäre und der Teststimulus zueinander passen oder nicht (match vs. non-match) können verschiedene Stimuluspaarungen dann unterschiedliche Motorantworten assoziativ einleiten und Alternativen inhibitorisch unterdrücken.

Das delayed-match-to-sample Paradigma und seine Varianten gehen zwar über einfache Reiz-Reaktions-Schemata hinaus, insofern multimodale Integration und Gedächtnis zeitlicher Relationen eine Rolle spielen; es ist aber immer noch recht einfach verglichen mit den komplexen Erfordernissen und Möglichkeiten von Verhalten in natürlichen Umgebungen. Zudem ist es möglich, daß einfache Reiz-Reaktions-Schemata auch ohne Beteiligung des präfrontalen Cortex auskommen könnten (Lopes da Silva et al., 1990; Taylor, 1999). Aus dem bisher Gesagten können jedoch einige abstrakte Funktionsprinzipien abgeleitet werden, die auch bei der Erfüllung komplizierterer Aufgaben durch cortikale neuronale Netze beteiligt sein sollten. Ganz zentral sind hier zwei Eigenschaften: einerseits die Fähigkeit zur Repräsentation und kurzzeitigen Speicherung von Verhaltensereignissen, andererseits die Möglichkeit auf diesen Repräsentationen Operationen auszuführen, sie zu kombinieren und zu verändern. Wie beides zusammen zu praktisch beliebig komplexem Verhalten führen kann, wollen wir abschließend darlegen.

Ausgehen wollen wir hierbei wieder von der Annahme, daß Verhaltensereignisse in lokalen Assemblies, z.B. präfrontal, in Form rekurrenter Aktivität gespeichert werden. Diese lokalen Netzwerke enthalten vermutlich wenig Information über den Stimulus selbst, nämlich kaum mehr, als das er aufgetreten ist oder nicht; ihre Aktivierung stellt gewissermaßen logisch-symbolische Variable für Ereignisse in einer Verhaltensaufgabe dar (vgl. Fuster, 1994; Platt und Glimcher, 1999). Die Art der Ereignisse definiert sich über Wechselwirkungen mit globalen Assemblies in anderen Cortexarealen. In Erweiterung der Verhältnisse beim delayed-matching-to-sample-Paradigma müssen diese Speichersemblies jedoch nicht repräsentativ für ganz konkrete Stimulus- oder Verhaltensmuster stehen. Sie können in gleicher Weise abstrakte innere Ereignisse repräsentieren, derart daß eine geeignete Menge solcher Assemblies den momentanen Zustand bei der Ausführung einer Aufgabe charakterisiert. Dies schließt ein, daß Speichersemblies selbst wieder miteinander kombiniert werden können, beispielsweise, wenn es bei eine Aufgabe auf die zeitliche Reihenfolge zweier Ereignisse A und B ankommt. Die Reihenfolge des Auftretens kann selbst wieder durch zwei lokale Assemblies AB und BA dargestellt werden.

Neben dieser Zustandscharakterisierung benötigt man zudem noch Interaktionen zwischen Assemblies, genauer gesagt einen Mechanismus, der auf ihnen Operationen ausführen kann, sie kombiniert, mit einlaufenden Sinnesdaten in Zusammenhang bringt,

und die zeitliche Ordnung von Ereignissen koordiniert, nun speziell auch durch die interne Detektion und Speicherung innerer Zustände, die den bisherigen Ablauf einer Handlung charakterisieren. Dieser Mechanismus repräsentiert relationales, temporales und logisches Wissen über eine Aufgabe. Aus Sicht der Informatik kann man sagen, es handelt sich dabei um die Implementierung eines Algorithmus, also um ein Programm.

Um Netzwerke zu konstruieren, die derartiges prozedurales Wissen implementieren, reicht interessanter Weise ein einziges weiteres Grundprinzip neben der Annahme von lokalen Speicherensembles und der gerichteten Aktivitätsausbreitung in Form sequentieller Aktivierung von Subensembles in Synfire-chains. Dies ist die Möglichkeit konditionierter Übergänge zwischen Speicherensembles in Abhängigkeit von der Aktivität in anderen Teilen des Netzwerks (Wennekers, 1998, 1999; Wennekers und Palm, 2000). Aktivität in einem Assembly kann sich dann abhängig von einem Eingangssignal auf ein anderes Assembly übertragen. Die Auslösung der Motoraktion durch ein präfrontales Motor-Set-Assembly und das plötzliche Eintreffen des Teststimulus ist ein Beispiel hierfür: der Teststimulus bzw. das entsprechende ihn darstellende globale Assembly ist das Eingangssignal, welches zur Übertragung der Aktivität vom Motor-Set-Assembly in höhere motorische Areale führt. Je nachdem ob der Teststimulus zum ersten Reiz paßt oder nicht, kann unterschiedliches Verhalten eingeleitet werden, d.h. der Übergang ist eingabeabhängig. Die genauen Vorgänge mögen komplizierter sein, das abstrahierte Grundprinzip eingabeabhängiger Übergänge erweist sich jedoch als fundamental zur Erklärung komplizierter Verhaltensmuster. Weiter oben hatten wir lineare Ketten assoziativ gespeicherter Muster als einfache Mechanismen für Zeitmustererkenner und -generatoren eingeführt. Eingabeabhängige Übergänge erweitern diese linearen Ketten nun auf praktisch beliebig komplexe Graphen, deren Knoten Subensembles entsprechen und deren Kanten mögliche Übergänge zwischen Knoten angeben, die aber im allgemeinen nur in Gegenwart einer gleichzeitigen passenden Eingabe wirklich ausgeführt werden können, d.h. die Übergänge sind konditioniert. Netzwerke von diesem Typ haben wir Synfire Graphen genannt in Anlehnung an den von Abeles geprägten Begriff der Synfire Chain für die deterministische Ausbreitung neuronaler Aktivität entlang linearer Ketten von Subensembles (Abeles, 1991; Wennekers, 1998, 1999). Synfire Graphen repräsentieren komplexes Zeitverhalten und das Zusammenspiel zwischen internen Zuständen und externen Stimuli. Man kann das Konzept zudem genauer formalisieren und nachweisen, daß Synfire Graphen beliebige deterministische und nichtdeterministische endliche Automaten implementieren können. Daß heißt, Synfire-Graph-Netzwerke sind in der Lage reguläre Sprachen zu generieren und zu erkennen (Hopcroft und Ullman, 1979; Wennekers, 1998, 1999).

Abbildung 8 zeigt schematisch noch einmal die Grundprinzipien: A und B stellen die Grundbausteine dar, nämlich (A) die gerichtete Aktivitätsausbreitung im Sequenzspeicher und (B) die Möglichkeit der Kurzzeitspeicherung in reverberierenden Assemblies. Durch eine Menge solcher Speicherensembles wie in B kann der momentane Zustand bei der Ausführung einer Verhaltensaufgabe beschrieben werden. Zudem kann aus Sequenzstücken wie in A ein Synfire Graph zusammengesetzt werden, der prozedurales

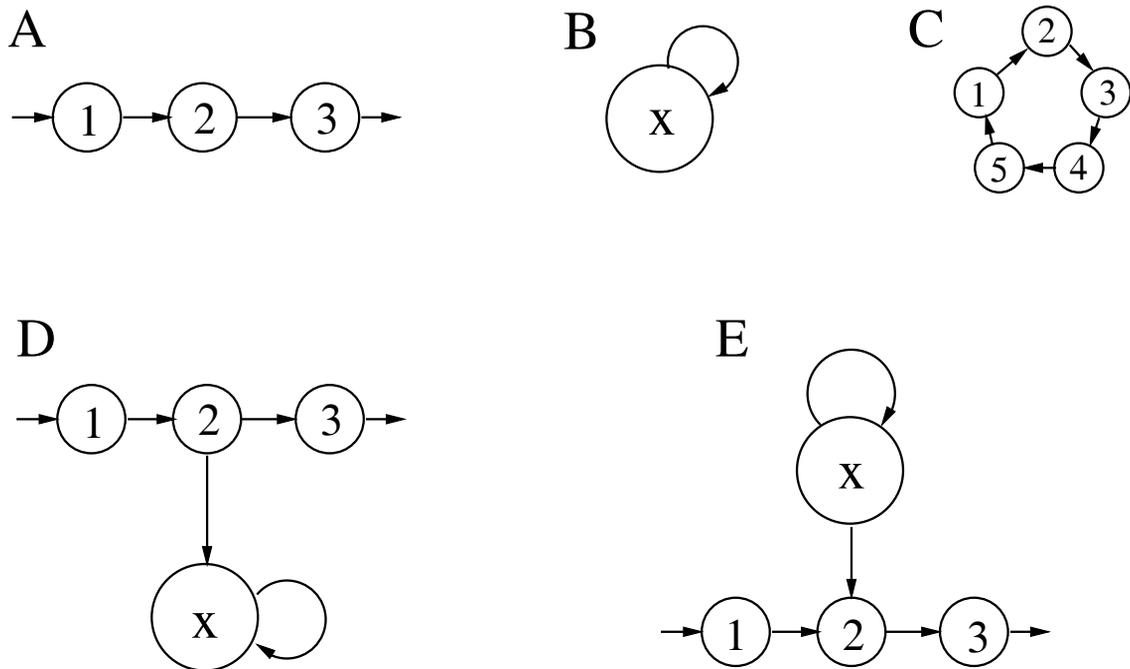


Abbildung 8: Elementare Bausteine für ein operationelles Assemblymodell. A zeigt exemplarisch ein (kurzes) Stück einer Mustersequenz. Hieraus lassen sich komplexe Graphen (Synfire-Graphen) zusammensetzen, die prozedurales, logisches oder zeitliches Wissen über eine Aufgabe repräsentieren können (Programmnetzwerke). In B ist ein Speicherassembly dargestellt; der rückkoppelnde Pfeil deutet an, daß es durch rekurrente Verbindungen in dem Assembly zum kurzzeitigen Halten der Netzwerkaktivität kommen kann. Abbildung C demonstriert, daß Speicherassemblies selbst aus rückgekoppelten Sequenzen bestehen können, daß also kein prinzipieller Unterschied bei der Implementierung von Speicher- und Programmnetzwerken besteht. D und E zeigen Wechselwirkungen zwischen Speicher- und Programmnetzwerk. D: Die gerichtete Aktivierung der Sequenz 1,2,3 erregt das Assembly x, welches anschließend signalisiert, daß die Sequenz bereits ausgeführt wurde. E: Falls die Schwelle des Knotens 2 relativ hoch ist, ist nur dann eine stabile Sequenzaktivierung möglich, wenn beim Feuern von Knoten 1 gleichzeitig auch das Speicherassembly x aktiv ist (konditionale Ausführung von Teilen des Synfiregraphen).

Wissen repräsentiert (Programmnetzwerk). Abbildung 8 C zeigt zudem, das Speicher-assemblys wie in B selbst wieder durch intern rückgekoppelte Sequenzen realisiert werden können, d.h. die Unterscheidung zweier Grundbausteine ist eine rein logische; prinzipiell können beide Prinzipien problemlos im selben Netzwerk implementiert werden. Hierbei ist insbesondere zu bemerken, daß Einzelzellen aufgrund des möglichen Überlapps in vielen Assemblys (Knoten bzw. Kreisen in Abbildung 8) enthalten sein können; daher können sie in der Regel kaum eindeutig dem Speicher- bzw. dem darauf operierenden Programmnetzwerk zugeordnet werden.

Die Abbildungen 8 D und E stellen zwei elementare Möglichkeiten der Interaktion zwischen dem Programm- und dem Speichernetzwerk dar. Zunächst können Programmknoten den Inhalt (i.e. den Aktivierungszustand) von Speicherknoten beeinflussen: wenn in D eine Aktivierungswelle die Knoten 1,2,3 entlang läuft, wird bei Erreichen des Knotens 2 das Speicherassembly x erregt. Seine Aktivierung signalisiert anschließend, daß der entsprechende Programmteil (z.B. korrespondierend mit einem Verhaltensereignis) bereits ausgeführt wurde. Ganz analog könnte der Knoten 2 auch direkt ein globales Motorassembly triggern, also ein Verhalten auslösen. Wenn x mit einem globalen sensorischen Assembly verbunden ist und seine Feuerschwelle recht hoch ist, kann das Feuern der Knotens 2 weiterhin erwirken, daß x in Abhängigkeit von der Aktivierung des globalen Assemblys erregt wird oder auch nicht (Schwellenkontrolle). Selbstverständlich können Verbindungen von Programm- zu Speicherassemblys auch hemmend sein, so daß die Aktivierung gespeicherter Ereignisse wieder unterdrückt werden kann, wenn sie nicht mehr notwendig ist.

Die Aktivität im Programmnetzwerk beeinflußt nicht nur die Aktivität in den Speicherassemblys, sondern der Fluß der Aktivität im Synfire-Graphen kann umgekehrt auch vom Zustand des Speichernetzwerks geregelt werden. Ein einfaches Beispiel ist in Abbildung 8 E zu sehen. Wenn die Aktivierungsschwelle des Knotens 2 hier relativ hoch ist, reicht die Aktivität von 1 nicht aus, um 2 zum Feuern zu bringen. Erst eine zusätzliche Erregung durch das Assembly x (wenn es aktiv ist) aktiviert auch 2, so daß eine stabile Sequenzausbreitung entlang der Kette 1,2,3 gewährleistet ist. D.h. der Aktivierungszustand des Assemblys x entscheidet über die konditionale Ausführung eines Teils des Synfire-Graphen. Hierdurch sind z.B. unterschiedliche Reaktionen in Verhaltensaufgaben möglich.

Es sollte klar sein, daß neben den einfachen Beispielen in Abbildung 8 A und B eine Vielzahl komplexerer Interaktionen zwischen Speicher- und Programmnetzwerken möglich ist. Prinzipiell können die Verhältnisse so verwickelt sein, daß in einem Netzwerk eine Unterscheidung zwischen beiden Komponenten gar nicht mehr eindeutig möglich ist. Die Unterscheidung ist aber zur Darstellung der Verarbeitungsprinzipien in Synfire-Graphen konzeptionell wichtig.

Abbildung 9 faßt nun unsere Auffassung von der Wirkungsweise operationeller Assemblys zusammen. Ausgehend von den über die Sinnesoberfläche einlaufenden Daten über die äußere Welt bilden sich in den sensorischen Hierarchien globale Assemblys als verteilte

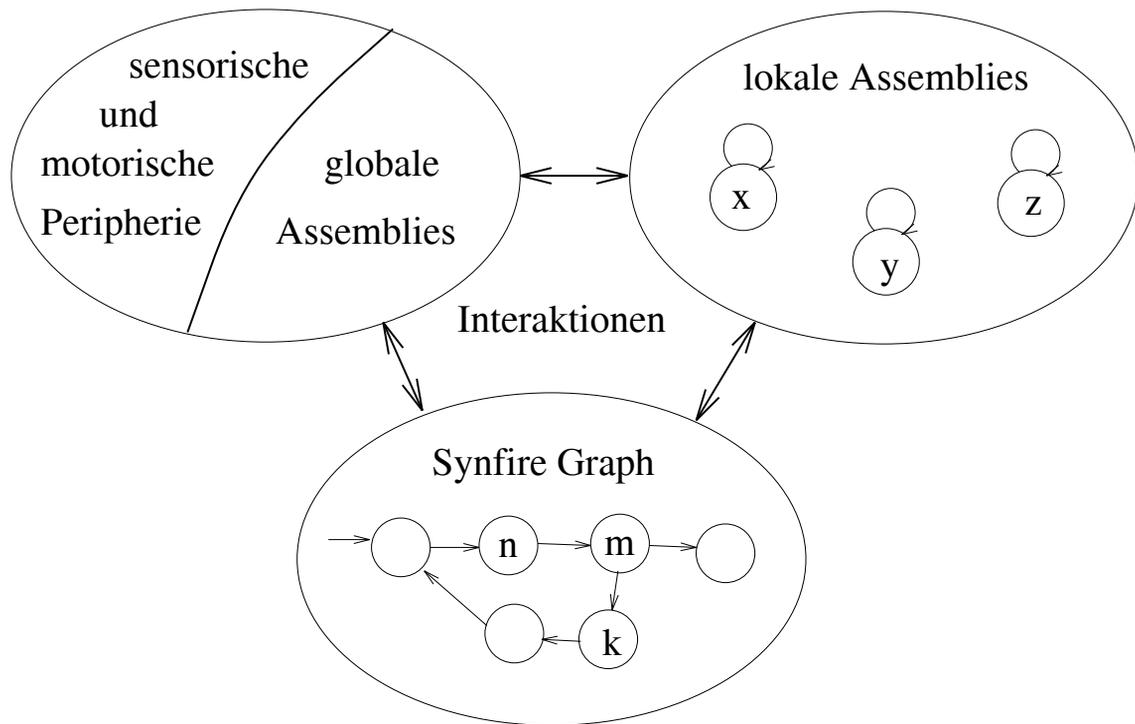


Abbildung 9: Einfaches Modell kognitiver Leistungen auf der Grundlage neuronaler Assemblies. Das Auftreten verhaltensrelevanter Ereignisse wird in lokalen Assemblies durch reverberierende Aktivität kurzzeitig gespeichert (x,y,z); Details repräsentieren hierbei globale, weit verteilte Assemblies mit Anknüpfung an die Sensomotorik. Reaktionsschemata, Handlungsabläufe und -alternativen regelt ein graphartiges Netzwerk, das in wechselseitiger Abhängigkeit mit den Speicherelementen und gegebenenfalls noch weiteren Strukturen steht.

Repräsentationen externer Ereignisse (links oben). Umgekehrt werden Motorprogramme detailliert in verteilten, komplex untereinander und mit der Sensorik vernetzten Assemblies abgelegt (diese werden durch "Training" aktiviert und Hebb'sch gelernt). In hohen Hirnbereichen existieren zudem lokale Assemblies, die repräsentativ für verhaltensrelevante Ereignisse stehen: x, y, z im Bild rechts; die rückkoppelnden Pfeile deuten an, daß diese Assemblies Kurzzeitspeichereigenschaften durch rekurrente Aktivität aufweisen. Zudem sind sie wechselseitig mit den entsprechenden globalen Assemblies vernetzt. Die Aktivierung eines globalen Assemblies kann so das lokale Memoryassembly anregen; umgekehrt kann von einem aktivierten Memoryassembly auf Eigenschaften des repräsentierten Ereignisses geschlossen werden (ein Spezialfall hiervon ist besonders auch die Auslösung einer Motoraktion). Interaktionen zwischen den lokalen und globalen Assemblies können selbstverständlich assoziativ und fehlertolerant sein, sowie weiteren Schwellenmodulationsmechanismen unterliegen (vgl. Eigenschaften des Assoziativspeichers weiter oben). Komplexe raumzeitliche und logische Beziehungen zwischen Verhaltensereignissen in einer gegebenen Aufgabe werden weiterhin in graphartigen Netzwerken mit konditionierten Übergängen repräsentiert (im unteren Teil des Bildes dargestellt). Der Fluß neuronaler Aktivität in diesem Netzwerk hängt vom Aktivitätszustand der Speichersemblies ab: je nachdem wie gewisse lokale Assemblies aktiviert sind - also die entsprechenden Verhaltensereignisse eingetroffen sind oder nicht - können unterschiedliche Teile des Programmnetzwerks konditional aktiv werden. Umgekehrt kann das Feuern von Subsemblies in diesem Netzwerk auch Einfluß auf die Aktivität der Speichersemblies nehmen, beispielsweise indem bei Eintritt gewisser interner Ereignisse entsprechende Assemblies aktiviert oder deaktiviert werden (eine Deaktivierung der cue-cell-Aktivität ist etwa kurz nach der Einleitung der Motoraktion im delayed-matching-to-sample-Paradigma beobachtbar). Die Bedingungen, Reihenfolge, etc., unter denen dies geschieht hängt natürlich vom jeweils aktivierten Bereich des Graphen ab, also vom bisherigen Verlauf der Verhaltensaufgabe. Auf diese Weise lassen sich praktisch beliebig komplexe Aufgaben in neuronalen Netzen basierend auf Cell-Assemblies darstellen (Wennekers 1998, 1999; Wennekers und Palm, 2000). Die hierbei getroffene Trennung zwischen quasi deklarativen Speichersemblies und operativen Netzwerkkomponenten in Form von Synfire-Graphen ist allerdings sicherlich eine starke Vereinfachung. Vermutlich dürften deklarative und prozedurale Komponenten in wirklichen neuronalen Netzen oft kaum voneinander zu trennen sein.

Nach der bisher dargelegten Vorstellung ist sowohl das deklarative als auch das prozedurale Wissen in Assemblies und ihren Verbindungen untereinander in Synfire-Graphen im Cortex niedergelegt. Es wurde dort durch Hebb'sche synaptische Plastizität gelernt. Obwohl wir glauben, daß allein der Cortex (mit Hilfe entsprechender Schwellensteuerung) schon zu diesen Leistungen fähig ist, ist uns natürlich bewußt, daß der Cortex hierbei mit einer Vielzahl anderer Hirnregionen zusammenarbeitet. Diese Zusammenarbeit ist sicherlich für die meisten realen Verhaltenssituationen, insbesondere komplexe motorische Aufgaben oder Einbeziehung emotionaler Faktoren, äußerst wichtig. Sie wird auch in der Neurobiologie und Hirnforschung untersucht. Ihre Einbeziehung würde aber den Rahmen

dieses Artikels sprengen.

Wie bereits gesagt lassen sich die oben beschriebenen Mechanismen im Rahmen der theoretischen Informatik genauer formalisieren. Man kommt dann zu neuronalen Netzen, die "endliche Automaten" im Sinne der Theorie formaler Sprachen implementieren. Dies heißt natürlich nicht, daß wir etwa glauben, daß natürliche Sprachen besonders gut oder sinnvoll im Rahmen der Theorie formaler Sprachen verstanden werden können. Außerdem wäre eine Anwendung der hier entwickelten Vorstellungen auf die cortikale Verarbeitung natürlicher Sprache nur ein Spezialfall. Sprachen sind in der Informatik universelle Konzepte zur Beschreibung von strukturierter Information, Automaten dienen der Verarbeitung solcher Information, in erster Linie ihrer Generierung und Erkennung (Hopcroft und Ullman, 1979). Tatsächlich stellt die Theorie formaler Sprachen allgemeine Konzepte bereit, die auf jede Art informationsverarbeitender Maschinen anwendbar sind, gerade so, wie die Mathematik auf die physikalische Welt universell anwendbar ist. Selbstverständlich paßt nicht jede Formel zu jedem Phänomen und ebenso wie die Mathematik noch nicht für alle Phänomene der physikalischen Welt voll entwickelt ist, sind auch die Konzepte der (Neuro-)Informatik und Automatentheorie noch nicht völlig entwickelt im Hinblick auf Informationsverarbeitungsprozesse im Gehirn. Wir sind allerdings schon der Meinung, daß die hier entwickelte Anwendung dieser Konzepte über die simple Computermetapher hinausgehend sinnvoll ist, wenn es auf logische oder syntaktische Strukturen bei der cortikalen Informationsverarbeitung ankommt. Im Rahmen der Erzeugung und des Verstehens natürlicher Sprache beispielsweise sind probabilistische Automaten (Hidden Markov Modelle) beliebte, plausible, wenn auch einfache und zum Teil umstrittene, Beschreibungsmittel, die sehr wahrscheinlich auch auf neuronaler Ebene implementiert sind, zweifelsohne aber im Detail komplexer als das abstrakte Konzept es vorsieht. Ganz vergleichbar strukturiert Verhaltenskontrolle handlungsrelevante Ereignisse in der Zeit (sozusagen Verhaltenssilben). Im oben beschriebenen Modell entwickeln wir Vorstellungen aus klassischen Ansätzen der Assemblytheorie heraus, welche die Assemblytheorie um operationelle Komponenten erweitern und zu einem Verständnis von komplexem Verhalten beitragen können. Formal sind diese operationellen Komponenten im Sinne der Automatentheorie beschreibbar. Die globale Assemblie Theorie zeigt hier Mechanismen und Strukturen auf, die ohne weiteres zur Realisierung solcher endlicher Automaten ausreichen. Dies ist gewissermaßen das Minimum, was man von jeder neurokognitiven Theorie erwarten sollte. Die Logik der Verhaltensaufgabe legt hier die Syntax fest und spiegelt sich in der Verbindungsstruktur des Synfire Graphen wieder; die Beziehung zwischen externen Ereignissen in der äußeren Welt und ihren internen Repräsentationen im Gehirn bestimmt zudem die Semantik des neuronalen Automaten. Durch die Idee der verteilten Repräsentationen gewinnt das hier entwickelte Automatenbild eine interessante zusätzliche Dimension: Die Beziehungen und Übergänge zwischen den verschiedenen Zuständen sind interessanter und reichhaltiger als bei den rein symbolisch arbeitenden Automaten der Informatik. Durch die verteilte Repräsentation ist ein mehr oder weniger großer Überlapp zwischen Konzepten möglich und damit eine nahezu

kontinuierliche Darstellung von Ähnlichkeiten (vgl. Abb. 1). Dies erweitert das schlichte Computerbild in eine Richtung, die von der praktischen Informatik heute in sogenannten hybriden Systemen verfolgt wird, in der subsymbolische und symbolische Informationsverarbeitungsparadigmen sinnvoll verbunden werden (Palm et al., 1991; Mikkulainen, 1993; Palm und Kraetzschmar, 1997)

Die Computermetapher ist in der Hirntheorie nicht neu. McCulloch und Pitts haben bereits 1943 gezeigt, daß einfache Netzwerke aus Schwellenneuronen im Prinzip jede aussagenlogische Formel repräsentieren und berechnen können und auch die syntaktische Korrektheit von Ausdrücken aus regulären Sprachen erkennen können (siehe auch Kleene, 1956). Interessanterweise hat sich auch von Neumann bei der Entwicklung des heute immer noch in den meisten Computern vorzufindenden Architekturkonzepts von frühen Vorstellungen über die physikalische Funktionsweise des Gehirns leiten lassen (von Neumann, 1958). Sogar das in der Informatik fundamentale theoretische Konzept der Turingmaschine entstammt einer Formalisierung des Denkprozesses; es soll die introspektive zugänglichen psychologischen Vorgänge beim Lösen mathematischer Gleichungen nachbilden. Neuere Ansätze betrachten zudem die Berechnungsstärke von Netzwerken spikender Neurone, parallele Algorithmen und Sprachen, berechnungsstärkere Maschinen als endliche Automaten (i.e. Kellerautomaten oder Turingmaschinen), probabilistische Automaten (Hidden Markov Modelle u.ä.), das Lernen von Automaten aus Beispielen, usw. (Beispiele in Shastri und Ajjanagadde, 1993; Maass, 1995; Omlin und Giles, 1996; Wennekers, 1998; Omori et al., 1999). Alle diese Ansätze lassen sich im Hinblick auf kognitive Prozesse und repräsentationale Strukturen im Gehirn diskutieren. In der vorliegenden Arbeit haben wir jedoch versucht in allgemeinerer Weise einen konzeptionellen Rahmen zu entwickeln, wie sich kognitive Leistungen im Sinne der Assemblytheory interpretieren lassen. Hierdurch wird ein direkterer Bezug zu den unterliegenden physiologischen Prozessen hergestellt. Zudem erscheint die Informationsverarbeitung im Gehirn in einem homogeneren Licht, als bei der Betrachtung von isolierten Einzelleistungen. Natürlich kann unser Entwurf nicht viel mehr als eine grobe Skizze sein; ein vollständiges Bild von der Rolle des Einzelneurons bis hin zu funktionellen Systemen wird sicherlich noch auf unabsehbare Zeit das kollektive Zusammenwirken von Forschern in allen Bereichen der Kognitionswissenschaften erfordern.

Literatur

- [1] Abbott, L.F.; Blum KI (1996) Functional significance of long-term potentiation for sequence learning and prediction. *Cerebral Cortex* 6, 406-416.
- [2] Abbott, L.; Varela, K.; Sen, K.; Nelson, S.B. (1997) Synaptic depression and cortical gain control. *Science* 275, 220-223.
- [3] Abeles, M. (1991) *Corticonics: Neural circuits of the cerebral cortex*. Cambridge University Press, Cambridge UK.

- [4] Abeles, M., Bergman, H. Margalit, E., and Vaadia, E. (1993) Spatio-temporal firing patterns in frontal cortex of behaving mokeys. *J.of Neurophysiol.* 70, 1629-1643.
- [5] Aertsen, A. (ed.) (1993) *Brain Theory*, Elsevier Science Publishers, Amsterdam,
- [6] Aertsen, A.; Arndt, M. (1993) Response synchronization in the visual cortex. *Current Opinion in Neurobiology* 3, 586-594.
- [7] Aertsen, A.M.H.J.; Gerstein, G.L.; Habib, M.K.; Palm, G. (1989) Dynamics of Neuronal Firing Correlation: Modulation of "Effective Connectivity". *J. Neurophysiol.* 61, 900-917.
- [8] Aertsen, A.; Erb, M.; Palm, G. (1994) Dynamics of Functional Coupling in the Cerebral Cortex: An Attempt at a Model-Based Interpretation. *Physica D* 75, 103-128.
- [9] Aertsen, A.; Erb, M.; Palm, G.; Schüz, A. (1994) Coherent Assembly Dynamics in the Cortex: Multi-Neuron Recordings, Network Simulations and Anatomical Cosiderations. In: *Proceedings Symposium "Oscillatory Event Related Brain Dynamics"*, Plenum.
- [10] Albus, J.S. (1975) A new approach to manipulator control: The cerebellar model articulation controller (CMAC). In: *Transactions of the ASME*, September, 220-227.
- [11] Albus, J.S. (1991) Outline for a theory of intelligence. *IEEE Transactions on Systems, Man, and Cybernetics* 21(3), 473-509.
- [12] Albus, J.S. (1992) RCS: A reference model architecture for intelligent control. *IEEE Special Issue of "Computer"*, entitled computer architectures for intelligent machines.
- [13] Arbib, M.A. (ed) (1994) *The Handbook of Brain Theory and Neural Networks*. MIT Press, Cambridge, MA.
- [14] Artola, A.; Singer, W. (1993) Long-term depression of excitatory synaptic transmission and its relationship to long-term potentiation. *Trends in Neurosciences* 16, 480-487.
- [15] Amari, S.I.; Joshida, K.; Kanatani K.I. (1977) A mathematical foundation for statistical neurodynamics. *SIAM J. Appl. Math.* 33, 95-126.
- [16] Amit, D.J. (1988) *Modeling Brain Function*. Cambridge University Press.
- [17] Amit, D.J. (1995) The Hebbian Paradigm Reintegrated: Local Reverberations as Internal Representations. *Behavioral and Brain Sciences* 18, 617-657.
- [18] Bibbig, A.; Wennekers, T.; Palm, G. (1995) A Neural Network Model of the Cortico-Interplay and the Representation of Contexts. *Behavioural Brain Research* 66, 169-175.
- [19] Bibbig, A., Wennekers, Th. (1996a) Hippocampal two-stage Learning and Memory Consolidation. In: *Proceedings of the 13th European Meeting on Cybernetics and Systems Research*, Vienna.

- [20] Bibbig, A.; Wennekers, Th.; Palm, G. (1996b) A neural network model of the cortico-hippocampal interplay: Contexts and generalization. *Engineering Applications of Artificial Intelligence*, 9 (2), 145-151.
- [21] Bienenstock, E. (1995) A model of the neocortex. *Network* 6, 179-224.
- [22] Bienenstock, E.L.; Cooper, L.N.; Munro, P.W. (1982) Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex. *Journal of Neuroscience* 2, 32-48.
- [23] Bishop, C.M. (1995) *Neural networks for pattern recognition*. Clarendon Press, Oxford.
- [24] Bliss, T.V.P.; Collingridge, G.L. (1993) A synaptic model of memory: long-term potentiation in the hippocampus. *Nature* 361, 31-39.
- [25] Blum, K.I.; Abbott, L.F. (1996) A model of spatial map formation in the hippocampus of the rat. *Neural Computation* 8, 85-93.
- [26] Bol, G.; Nakhaeizadeh, G.; Vollmer, K.H. (1995) *Finanzmarktanalyse und -prognose mit innovativen Verfahren*. Physica Verlag.
- [27] Braitenberg, V. (1978) *Cell Assemblies in the Cerebral Cortex*. In: Heim, R., Palm, G. (eds.) *Theoretical Approaches to Complex Systems*, Springer, Berlin, 171-188.
- [28] Braitenberg, V.; Schüz, A. (1991) *Anatomy of the cortex*. Springer, Berlin.
- [29] Braitenberg, V.; Pulvermüller, F. (1992) Entwurf einer neurologischen Theorie der Sprache. *Naturwissenschaften* 79, 103-117.
- [30] Brause, R. (1995) *Neuronale Netze*. Teubner Verlag, Stuttgart.
- [31] Brown, M.A.; Sharp, P.E. (1995) Simulation of spatial learning in the Morris water maze by a neural network model of the hippocampal formation and nucleus accumbens. *Hippocampus* 5, 171-188.
- [32] Burgess, N.; Donnett, J.G.; Jeffery, K.J.; O'Keefe, J. (1997) Robotic and neuronal simulation of the hippocampus and rat navigation. *Philosophical Transactions of the Royal Society of London B* 352, 1535-1543.
- [33] Bullier, J.; Munk, H.J.; Nowak, L.G. (1993) Corticocortical connections sustain interarea synchronization. *Concepts Neurosci.* 4, 159-174.
- [34] Burdet, E.; Milner, T.E. (1998) Quantization of human motions and learning of accurate movements. *Biological Cybernetics* 78, 307-318.
- [35] Carpenter, G.A.; Grossberg, S. (1991) *Pattern recognition by self-organizing neural networks*. MIT press, Cambridge, MA.

- [36] Cohen, M.A.; Grossberg, S. (1986) Neural dynamics of speech and language coding: Developmental programs, perceptual grouping, and competition for short term memory. *Human Neurobiology* 5, 1-22.
- [37] Damasio, A.R. (1989) Time-locked multiregional retroactivation: A systems-level proposal for the neural substrates of recall and recognition. *Cognition* 33, 25-62.
- [38] Damasio H.; Grabowski, T.J.; Tranel, D.; Hichwa, R.D.; Damasio, A.R. (1996) A neural basis for lexical retrieval. *Nature* 380, 499-505.
- [39] Desimone, R.D.; Duncan, J. (1995) Neural Mechanisms of Selective Visual Attention. *Annual Review of Neuroscience* 18, 139-222.
- [40] Doya, K. (1999) What are the computations of the cerebellum, the basal ganglia and the cerebral cortex? *Neural Networks* 12, 961-974.
- [41] Eckhorn, R.; Bauer, R.; Jordan, W.; Brosch, M.; Kruse, W.; Munk, M.; Reitboeck, H.J. (1988) Coherent oscillations: A mechanism of feature linking in the visual cortex? *Biol.Cyb.* 60, 121-
- [42] Eckhorn, R. (1999) Neural mechanisms of scene segmentation: Recordings from the visual cortex suggest basic circuits for linking field models. *IEEE Transactions on Neural Networks* 10, 464-479.
- [43] Edelman, G.M. (1992). *Bright air, brilliant fire: on the matter of the mind.* Basic Books, New York.
- [44] Elman, J.L. (1990) Finding structure in time. *Cognitive Science* 14, 179-211.
- [45] Erb, M.; Aertsen, A. (1992) Dynamics of activity in biology-oriented neural network models: stability at low firing rates. In: Aertsen, A. and Braitenberg, V. (eds.) *Information Processing in the Cortex. Experiments and Theory.* Springer, Berlin.
- [46] Erwin, E.; Obermayer, K.; Schulten, K. (1995) Models of orientation and ocular dominance columns in the visual cortex: a critical comparison. *Neural Computation* 7, 425-486.
- [47] Feldman, J.A.; Ballard, D.H. (1982) Connectionist models and their properties. *Cognitive Science* 6, 205-254.
- [48] Felleman, D.; van Essen, D. (1991) Distributed Hierarchical Processing in the Primate Cerebral Cortex. *Cerebral Cortex* 1, 1-47.
- [49] Franz, M.O.; Mallot, H.P. (2000) Biomimetic robot navigation. *Robotics and Autonomous Systems* 30, 133-153.
- [50] Fukushima, K.; Myaki, S. (1982) Neocognitron: A new algorithm for pattern recognition tolerant of deformations and shifts in position. *Pattern Recognition* 15, 455-469.

- [51] Fuster, J.M. (1973) Behavioural electrophysiology of the prefrontal cortex. *J. Neurophysiol.* 36, 61-78.
- [52] Fuster, J.M. and Jervey J.P. (1980) Inferotemporal neurons distinguish and retain behaviorally relevant features of visual stimuli. *Science* 212, 952-955.
- [53] Fuster, J.M. (1989) *The prefrontal cortex: anatomy, physiology, and neuropsychology of the frontal lobe.* New York: Raven Press.
- [54] Fuster, J.M. (1994) *Memory in the cerebral cortex. An empirical approach to neural networks in the human and nonhuman primate.* Cambridge, MA: MIT Press.
- [55] Georgopolous, A.P. (1986) On reaching. *Annual Review of Neuroscience* 9, 147-170.
- [56] Gerstein, G.L.; Bedenbaugh, P.; Aertsen, A.M.H.J. (1989) Neuronal assemblies. *IEEE Transactions on Biomedical Engineering* 36, 4-14.
- [57] Gerstner, W. and v.Hemmen, J.L. (1993a) Associative memory in a network of 'spiking' neurons. *Network* 3, 139-64.
- [58] Gerstner, W. (1995) Time Structure of the Activity in Neural Network Models. *Phys.Rev.E* 51, 738-758.
- [59] Gerstner, W., Ritz, R., van Hemmen, J.L. (1993b) A biologically motivated and analytically soluble model of collective oscillations in the cortex, I. Theory of weak locking. *Biol.Cybern.* 68, 363-374.
- [60] Gerstner, W.; Abbott, L.F. (1997) Learning navigational maps through potentiation and modulation of hippocampal place cells. *J.Comp.Neurosci.* 4, 79-94.
- [61] Gray, C.M. (1994) Synchronous Oscillations in Neuronal Systems: Mechanisms and Functions. *J.Comp. Neurosci.* 1, 11-38.
- [62] Gross, H.M.; Heinze, A; Seiler, T.; Stephan, V. (1999) Generative character of perception: a neural architecture for sensorimotor anticipation. *Neural Networks* 12, 1101-1121.
- [63] Grossberg, S. (1999b) How does the cerebral cortex work? Learning, attention, and grouping by the laminar circuits of visual cortex. *Spatial Vision* 12, 163-186.
- [64] Grossberg, S. (1999b) The link between attention, brain learning, and consciousness. *Consciousness and Cognition* 8, 1-44.
- [65] Harris, K.D.; Recce, M. (1997) Absolute localization for a mobile robot using place cells. *Robotics and autonomous systems* 22, 393-406.
- [66] Harth, E.M.; Csermely T.J.; Beek, B. and Lindsay, R.D. (1970) Brain Functions and Neural Dynamics. *J.Theoret.Biol.* 26, 93-120.

- [67] Haykin, S. (1994) Neural networks: a comprehensive foundation. Macmillan College Publishing Company.
- [68] Hebb, D.O. (1949) The organization of behavior. A neuropsychological theory. Wiley, New York.
- [69] Hetherington, P.A.; Shapiro, M.L. (1993) Simulating Hebb cell assemblies: the necessity for partitioned dendritic trees and a post-not-pre LTD rule. *Network* 4, 135-153.
- [70] Herrmann, M., Hertz, J.A., and Prügel-Bennett, A. (1995) Analysis of Synfire Chains. *Network* 6, 403-14.
- [71] Hopcroft, J.E.; Ullman, J.D. (1979) Introduction to Automata Theory, Languages, and Computation. Addison-Wesley, Reading, MA.
- [72] Hopfield, J.J. (1982) Neural Networks and Physical Systems with Emergent Collective Computational Properties. *Proc.Natl.Acad.Sci USA* 79, 2554-2558.
- [73] Hubel, D.H.; Wiesel, T.N. (1962) Receptive fields, binocular interaction and functional architecture in the cats visual cortex. *J.Physiol.Lond.* 160, 106-154.
- [74] Huening, H.; Gluender, H.; Palm, G. (1998) Synaptic delay learning in pulse-coupled neurons. *Neural Computation* 10, 555-565.
- [75] Itti, L.; Koch, C.; Niebur, E. (1998) A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis* 20(11), 1254-1259.
- [76] Iwamura, Y. (1998) Hierarchical somatosensory processing. *Current Opinion in Neurobiology* 8, 522-528.
- [77] Jensen, O. and Lisman, J.E. (1996) Novel Lists of 7 plus-minus 2 Known Items Can Be Reliably Stored in an Oscillatory Short-Term Memory Network: Interaction with Long-Term Memory. *Learning and Memory* 3, 257-263.
- [78] Kirkland, K.L.; Gerstein, L.G. (1999) A Feedback Model of Attention and Context Dependence in Visual Cortical Networks. *J.Comp.Neuroscience*, in press.
- [79] Kleene, S.C. (1956) Representation of Events in Nerve Nets and Finite Automata. In C.E. Shannon and J.McCarthy (eds.) *Automata Studies*. Princeton Press, Princeton.
- [80] Klimesch, W. (1994) The structure of long-term memory: A connectivity model of semantic processing. Hillsdale, NJ: Lawrence Erlbaum.
- [81] Koch, C.; Davis, J.L. (eds) (1994) *Large Scale Neuronal Theories of the Brain*. MIT Press, Cambridge, MA.
- [82] Koch, C.; Segev, I. (eds.) (1989) *Methods in Neural Modeling: from Synapses to Networks*. MIT Press, Cambridge, MA.

- [83] Koch, C.; Ullman, S. (1985) Shifts in selective visual attention: towards the underlying neural circuitry. *Human Neurobiology* 4, 219-227.
- [84] Lankheet, M.J.M., Molenaar, J., and van de Grind, W.A. (1989) The Spike Generating Mechanism of Retinal Ganglion Cells. *Vision Research* 29, 505-517.
- [85] Lansner, A.; Fransen, E. (1992) Modelling Hebbian cell assemblies comprised of cortical neurons. *Network* 3, 105-119.
- [86] Lashley, K.S. (1950) In search of the engram. *Symposium of the Society for Experimental Biology* 4, 454-482.
- [87] Lieblein, I.; Arbib, M.A. (1982) Multiple representations of space underlying behavior. *Behavioral and Brain Sciences* 5, 627-659.
- [88] Lippmann, R.P. (1989) Review of neural networks for speech recognition. *Neural Computation* 1, 1-38.
- [89] Little, W.A. (1974) The existence of persistent states in the brain. *Math. BioSci.* 19, 101-120.
- [90] Livingstone, M.S.; Hubel, D.H. (1987) Psychophysical Evidence for Separate Channels for the Perception of Form, Color, Movement and Depth. *Journal of Neuroscience* 7(11), 3416-3468.
- [91] Livingstone, M.; Hubel, D.H. (1988) Segregation of form, color, movement and depth: anatomy, physiology, and perception. *Science* 240, 740-49.
- [92] Logothetis, N.K.; Sheinberg, D.L. (1996) Visual object recognition. *Annu.Rev.Neurosci.* 19, 577-621.
- [93] Lopes da Silva, F.H.; Witter, M.P.; Boeijinga, P.H.; Lohmann, A.H.M. (1990) Anatomic Origin and Physiology of the Limbic Cortex. *Physiological Reviews* 70, 453-511.
- [94] Maass, W. (1995) Lower bounds on the Computational Power of Networks of Noisy Spiking Neurons. *Neural Computation* 8, 1-20.
- [95] Magee, J.C.; Johnston, D. (1997) A Synaptically Controlled, Associative Signal for Hebbian Plasticity in Hippocampal Neurons. *Science* 275, 209-213
- [96] Mallot, H.A. (1997) Behavior-oriented approaches to cognition: Theoretical perspectives. *Theory in Biosciences* 116, 192-220.
- [97] von der Malsburg, C. (1981) The correlation theory of brain function. Internal Report 81-2. Max-Planck Institut für Biophysikalische Chemie, Göttingen.
- [98] Markram, H.; Tsodyks, M. (1996) Redistribution of synaptic efficacy between neocortical pyramidal neurons. *Nature* 382, 807-810.

- [99] Martin, A.; Wiggs, C.L.; Ungerleider, L.G.; Haxby, J.V. (1996) Neural correlates of category-specific knowledge. *Nature* 379, 649-652.
- [100] McClelland, J.L.; Rumelhart, D.E.; Hinton, G.E. (1986) The appeal of parallel distributed processing. In: Rumelhart, D.E., McClelland, J.L. (eds.) *Parallel Distributed Processing*, Vol. 1. MIT Press, Cambridge, M.A., 3-44.
- [101] McClelland, J.L.; McNaughton, B.L.; O'Reilly, R.C. (1995) Why there are complementary Learning Systems in the Hippocampus and Neocortex: Insights from the Successes and Failures of Connections Models of Learning and Memory. *Psychological Review* 102, 419-457. McGregor, R.J. (1987) *Neural and Brain Modeling*. Academic Press, New York.
- [102] McCulloch, W.S.; Pitts, W. (1943) A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics* 5, 115-133.
- [103] McGregor, R.J. (1993) Composite cortical networks of multimodal oscillators. *Biol.Cybern.* 69, 243-255.
- [104] McQueen, J.M.; Cutler, A.; Briscoe, T.; Norris, D. (1995) Models of continuous speech recognition and the contents of the vocabulary. *Language and Cognitive Processes* 10(3), 309-331.
- [105] Merigan, W.H.; Maunsell, J.H.R. (1993) How Parallel are the Primate Visual Pathways? *Ann.Rev.Neurosci.* 16, 369-402.
- [106] Miikkulainen, R. (1993) *Subsymbolic natural language processing. An integrated model of scripts, lexicon, and memory*. MIT Press, Cambridge, Massachusetts.
- [107] Miller, R. (1991) *Cortico-Hippocampal Interplay and the Representation of Contexts in the Brain*. Springer, Berlin.
- [108] Miller, R. (Ed.) (2000) *Time and the Brain. Conceptual Advances in Brain Research*, vol.2. Harwood Academic Publishers (im Druck).
- [109] Milner, T.E. (1992) A model for the generation of movements requiring endpoint precision. *Neuroscience* 49, 365-374.
- [110] Milner, P.M. (1974) A model for visual shape recognition. *Psychological Review* 81(6), 521-535.
- [111] Milner, P.M. (1996) Neural Representations: Some Old Problems Revisited. *J.of Cognitive Neuroscience* 8, 69-77.
- [112] Miltner, W.H.; Braun, C.; Arnold, M.; Witte, H.; Taube, E. (1999) Coherence of gamma-band EEG activity as a basis for associative learning. *Nature* 397, 434-436.
- [113] Miyashita, Y.; Chang, H.S. (1988) Neural correlate of pictorial short term memory, *Nature* 331, 68-70.

- [114] Neumann, H.; Mössner, P. (1996) Neural Model of Cortical Dynamics in Resonant Boundary Detection and Grouping. Proc. ICANN-96, Bochum (Germany).
- [115] von Neumann, J. (1958) The computer and the brain. Yale University Press. New Haven, London.
- [116] Nunez, P.L. (1995) Neocortical Dynamics and Human EEG Rhythms. Oxford University Press, New York.
- [117] O'Keefe, J.; Nadel, L. (1978) The hippocampus as a cognitive map. Clarendon Press, Oxford.
- [118] Omlin, C.W.; Giles, C.L. (1996) Constructing Deterministic Finite-State Automata in Recurrent Neural Networks. Journal of the ACM 43(6), 937-972.
- [119] Omori, T.; Mochizuki, A.; Mizukati, A.; Nishizaki, M. (1999) Emergence of symbolic behaviour from brain like memory with dynamic attention. Neural Networks 12, 1157-1172.
- [120] Palm, G. (1980) On associative memories. Biol.Cybern. 36, 19-31.
- [121] Palm, G. (1982a) Neural Assemblies. An Alternative Approach to Artificial Intelligence. Springer, Berlin.
- [122] Palm, G. (1982b) Rules for Synaptic Changes and their Relevance for the Storage of Information in the Brain. In: Trappl, R. (ed.) Cybernetics and Systems Research. North-Holland Publishing Company.
- [123] Palm, G.; Aertsen, A.; eds. (1986) Brain Theory. Springer, Berlin, Heidelberg, New York, Tokyo.
- [124] Palm, G. (1986) Associative Networks and Cell Assemblies. In: Palm, G.; Aertsen, A. (eds.) Brain Theory. Springer.
- [125] Palm, G. (1990) Cell Assemblies as a Guideline for Brain Research. Concepts in Neuroscience 1, 133-147.
- [126] Palm, G.(1991) Memory Capacities of Local Rules for Synaptic Modification. Concepts in Neuroscience 2, 97-128.
- [127] Palm, G.; Rückert, U.; Ultsch, A. (1991) Wissensverarbeitung in neuronaler Architektur. In: Bauer, W.; Hernandez, D. (eds.) Informatik-Fachberichte 1991: Verteilte Künstliche Intelligenz und kooperatives Arbeiten. Springer, Berlin, Heidelberg, New York.
- [128] Palm, G. (1993a) On the internal structure of cell assemblies. In: Aertsen, A. (ed.) Brain Theory, Elsevier Science Publishers, Amsterdam, 261-271.

- [129] Palm, G. (1993b) Cell Assemblies, Coherence and Cortico-Hippocampal Interplay. *Hippocampus*, Vol. 3, 1 Special Issue (Nitsch, R., Ohm, T.G, eds.), 219-225.
- [130] Palm, G.; Sommer, F.T. (1996) Associative data storage and retrieval in neural networks. In *Models of neural networks III*, eds. Domany, E.; van Hemmen, J.L.; Schulten, K., 79-118. New York: Springer Verlag.
- [131] Palm, G.; Kraetzschmar, G. (1997) SFB 527: Integration symbolischer und subsymbolischer Informationsverarbeitung in adaptiven sensomotorischen Systemen. In: Jarke, M.; Pasedach, K.; Pohl, K. (eds.) *Informatik '97 - Informatik als Innovationsmotor*. Proceedings der 27. Jahrestagung der Gesellschaft für Informatik, Aachen, 1997, Springer, Berlin, Heidelberg.
- [132] Palm, G.; Wennekers, T. (1997) Synchronicity and its use in the brain. *Behavioral and Brain Sciences* 20, 695-696.
- [133] Pasemann, F. (1995) Neuromodules: A dynamical systems approach to brain modelling. In: Herrmann, H.J., Wolf, D.E., Pöppel, E. (eds.), *Supercomputing in Brain Research: From Tomography to Neural Networks*. World Scientific, Singapore, 1995, 331-348.
- [134] Perkel, D.H.; Gerstein, G.L.; Moore, G.P. (1970) Neuronal Spike Trains and Stochastic Point Processes. Part 1. The Single Spike Train. *Biophys. J.* 7, 391-418. Part 2. Simultaneous Spike Trains. *Biophys. J.* 7, 419-440.
- [135] Phillips, W.A.; Singer, W. (1997) In search of common foundations for cortical computation. *Behavioral and Brain Sciences* 20, 657-722.
- [136] Pinsky, P.F.; Rinzel, J. (1994) Intrinsic and Network Rhythmogenesis in a Reduced Traub Model for CA3 Neurons. *J.Comp.Neurosci.* 1, 39-60.
- [137] Platt, M.L.; Glimcher, P.W. (1999) Neural Correlates of Decision Variables in the Cerebral Cortex. *Nature* 400, 233-238.
- [138] Posner, M.I. (1990) The attention system of the human brain. *Ann.Rev.Neurosci.* 13, 25-42.
- [139] Pulvermüller, F.; Preißl, H. (1991) A cell assembly model of language. *Network* 2, 455-468.
- [140] Pulvermüller, F. (1992) Constituents of a neurological theory of language. *Concepts in Neuroscience* 3, 157-200.
- [141] Pulvermüller, F. (1998) On the matter of rules. Past-tense formation and its significance for cognitive neuroscience. *Network: Comput. Neural Syst.* 9, R1-R52.
- [142] Pulvermüller, F. (2000) Words in the brain's language. *Behavioural and Brain Sciences*, in press.

- [143] Rao, R.P.N.; Ballard, D.H. (1997) Dynamic model of visual recognition predicts neural response properties in the visual cortex. *Neural Computation* 9, 721-763.
- [144] Qin, Y.L.; McNaughton, B.L.; Skaggs, W.E.; Barnes, C.A. (1997) Memory reprocessing in corticocortical and hippocampocortical neuronal ensembles. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 352, 1525-1533.
- [145] Rauschecker, J.P. (1998) Cortical Processing of Complex Sounds. *Current Opinion in Neurobiology* 8, 516-521.
- [146] Recce, M.; Harris, K.D. (1996) Memory for places: A navigational model in support of Marr's theory of hippocampal function. *Hippocampus* 6, 735-748.
- [147] Renart, A.; Parga, N.; Rolls, E.T. (1999) Backward Projections in the Cerebral Cortex: Implications for Memory Storage. *Neural Computation* 11, 1349-1388.
- [148] Renart, A.; Parga, N.; Rolls, E.T. (1999) Associative memory properties of multiple cortical modules. *Network: Comutation in Neural Systems* 10, 237-255.
- [149] Ripley, B.D. (1996) Pattern recognition and neural networks. Cambridge University Press.
- [150] Robert, A. (1997) From Contour Integration to Image Schemas: A Modern Perspective on Ge- stalt Psychology. Technical Report 97.02, Dept. of Cognitive Science, University of California, San Diego.
- [151] Rodriguez, E.; George, N.; Lachaux, J.P.; Martinerie, J.; Renault, B.; Varela, F.J. (1999) Perception's shadow: long-distance synchronisation of human brain activity. *Nature* 397, 430-433.
- [152] Rojas, R. (1993) Theorie der neuronalen Netze. Springer Verlag.
- [153] Rumelhart, D.; McClelland, J. (1986a) PDP models and general issues in cognitve sciences. In: Rumelhart, D.; McClelland, J. (eds.) *Parallel distributed processing*. MIT Press, Cambridge, M.A.
- [154] Salinas, E.; Abbott, L.F. (1995) Transfer of coded information from sensory to motor networks. *J.Neurosci.* 15, 6461-6474.
- [155] Samsonovich, A.; McNaughton, B.L. (1997) Path integration and cognitive mapping in a continuous attractor neural network model. *J. Neurosci.* 17, 5900-5920.
- [156] Schillen, T.B.; König, P. (1994) Binding by temporal structure in multiple feature domains of an oscillatory neural network. *Biol.Cybern.* 70, 397-405.
- [157] Schöneburg, E. (1993) Industrielle Anwendungen neuronaler Netze. Addison Wesley.
- [158] Schürmann, J. (1996) Pattern classification - a unified view of statistical and neural approaches. John Wiley& Sons.

- [159] Schwenker, F.; Sommer, F.T.; Palm, G. (1996) Iterative Retrieval of Sparsely Coded Associative Memory Patterns. *Neural Networks*, Vol. 9, No. 3, 445-455.
- [160] Segev, I. (1992) Single neurone models: oversimple, complex, and reduced. *Trends in Neuroscience* 15, 414-421.
- [161] Sejnowski, T.J.; Rosenberg, C.R. (1987) Parallel networks that learn to pronounce. *Complex Systems* 1, 145-168.
- [162] Shastri, L.; Ajjanagadde, V. (1993) From simple associations to systematic reasoning: A connectionist representation of rules, variables and variable bindings. *Behavioral and Brain Sciences* 16, 417-494.
- [163] Shaw, G.; Harth, E; and Scheibel, A. (1982) Cooperativity in Brain Function: Assemblies of Approximately 30 Neurons. *Experimental Neurology* 77, 324-358.
- [164] Singer, W.; Gray, C.M. (1995) Visual feature integration and the temporal correlation hypotheses. *Ann.Rev.Neuroscience* 18, 555-586.
- [165] Sommer, F.T.; Palm, G. (1999) Improved Bidirectional Retrieval of Sparse Patterns Stored by Hebbian Learning. *Neural Networks* 12, 281-297.
- [166] Sommer, F.T.; Wennekers, Th.; Palm, G. (1998) Bidirectional completion of cell assemblies in the cortex. In: Bower, J.M. (ed.) *Computational Neuroscience: Trends in Research*. Plenum Press, New York.
- [167] Sommer, F.T.; Wennekers, T. (2000) Associative memory in a pair of cortical cell groups with reciprocal connections. Submitted.
- [168] Stöcker, M.; Reitboeck, H.J.; Eckhorn, R. (1996) A neural network for scene segmentation by temporal coding. *Neurocomputing* 11, 123-134.
- [169] Stein, R.B. (1967) Some Models of Neuronal Variability. *Biophysical Journal* 7, 37-68.
- [170] Sterratt, D.C. (1999) Is a biological temporal learning rule compatible with learning Synfire chains? *Proceedings of the 9th International Conference on Artificial Neural Networks, ICANN99*. IEE Conference Publication No.470, London, 551-556.
- [171] Suder, K.; Wörgötter, F.; (2000) The control of low-level information in the visual system. *Rev. in the Neurosciences* 11, 127-146.
- [172] Swindale, N.V. (1990) Is the cerebral cortex modular? *Trend in Neuroscience* 13, 487-492.
- [173] Tanaka, K. (1993) Neuronal Mechanisms of Object Recognition. *Science* 262, 685-688.
- [174] Tanaka, K. (1997) mechanisms of object recognition: monkey and human studies. *Current Opinion in Neurobiology* 7, 523-529.

- [175] Taylor, J.G. (1999) Towards the networks of the brain: from brain imaging to consciousness. *Neural Networks* 12, 943-959.
- [176] Tolle, H.; Ersü, E. (1992) Neurocontrol, learning control systems inspired by neuronal architectures and human problem solving strategies. *Lecture Notes in Control and Information Sciences*, No. 172.
- [177] Tononi, G.; Sporns, O.; Edelman, G.M. (1992) Reentry and the problem of integrating multiple cortical areas: simulation of dynamic integration in the visual system. *Cerebral Cortex* 2, 310-335.
- [178] Traub, R.D.; Wong, R.; Miles, R.; Michelson, H. (1991) A model of a CA3 hippocampal neuron incorporating voltage-clamp data on intrinsic conductances. *J.Neurophysiol.* 66:635-649.
- [179] Treisman, A. (1996) The binding problem. *Current Opinion in Neurobiology* 6, 171-178.
- [180] Tsumoto, T. (1992) Long-term potentiation and long-term depression in the neocortex. *Progress in Neurobiology* 39, 209-228.
- [181] Wennekers, T., Sommer, F.T., and Palm, G. (1995) Iterative Retrieval in Associative Memories by Threshold Control of Different Neural Models. In: Herrmann, H.J. and Wolf, D.E. and Pöppel, E. (eds.) *Supercomputing in Brain Research: From Tomography to Neural Networks*. World Scientific Publishing, Singapore, 301-319.
- [182] Wennekers, T. and Palm G. (1997) On the relation between neural modeling and experimental neuroscience. *Theory in Biosciences* 116, 273-289.
- [183] Wennekers, T. (1998) *Synfire Graphs: From Spike Patterns to Automata of Spiking Neurons*. Technical report Nr. 98-8, Fakultät für Informatik, Universität Ulm, Germany.
- [184] Wennekers, T. (1999) *Synchronisation und Assoziation in neuronalen Netzen*. Shaker Verlag, Aachen.
- [185] Wennekers, T.; Sommer, F.T. (1999) Gamma-oscillations support optimal retrieval in associative memories of two-compartment neurons. *Neurocomputing* 26-27, 573-578, 1999.
- [186] Wennekers, T. (2000) Dynamics of spatio-temporal patterns in associative networks of spiking neurons. *Neurocomputing*, in press.
- [187] Wennekers, T. and Palm G. (2000) Cell Assemblies, Associative Memory and Temporal Structure in Brain Signals. In: Miller, R. (ed.) *Time and the Brain. Conceptual Advances in Brain Research*, vol.2, Harwood Academic Publishers (im Druck).
- [188] Werbos, P.J. (1988) Backpropagation: past and future. In: *Proceedings of the IEEE International Conference on Neural Networks*, 343-353.

- [189] Wermter, S. (1995) Hybrid connectionist natural language processing. Chapman & Hall, London.
- [190] Wickelgren, W.A. (1969) Context-sensitive coding, associative memory, and serial order in (speech) behavior. *Psychological Review* 76, 1-15.
- [191] Wickelgren, W.A. (1977) *Learning and Memory*. Prentice-Hall, Englewood Cliffs, N.J.
- [192] Wickelgren, W.A. (1981) Human Learning and Memory. *Annual Review of Psychology* 32, 21-52.
- [193] Wickelgren, W.A. (1992) Webs, cell assemblies, and chunking in neural nets. *Concepts in Neuroscience* 3(1), 1-53.
- [194] Wickens, J.; Hyland, B.; Anson, G. (1994) Cortical cell assemblies: a possible mechanism for motor programs. *Journal of Motor Behavior* 26, 66-82.
- [195] Willshaw, D.J.; Buneman, O.P.; and Longuet-Higgins, H.C. (1969) Non-holographic associative memory. *Nature* 222, 960ff.
- [196] Wörgötter, F.; Koch, C. (1991) A Detailed Model of the Primary Visual Pathway in the Cat: Comparison of Afferent Excitatory and Intracortical Inhibitory Connection Schemes for Orientation Selectivity. *Journal of Neuroscience* 11, 1959-1979.
- [197] Wolpert, D.M. (1997) Computational approaches to motor control. *Trends in Cognitive Sciences* 1(6), 209-216.
- [198] Wright, J.J.; Liley, D.T.J. (1996) Dynamics of the brain at global and microscopic scales: Neural networks and the EEG. *Beh. Brain Sciences* 19, 285-320.
- [199] Young, M.P.; Scannell, J.W.; Burns, G. (1995) *The analysis of cortical connectivity*. Heidelberg: Springer.
- [200] Zhang, L.I.; Tao, H.W.; Holt, C.E.; Harris, W.A.; Poo, M. (1998) A critical window for cooperation and competition among developing retinotectal synapses. *Nature* 395, 37-44.
- [201] Zipser, D.; Kehoe, B.; Littlewort, G.; Fuster, J (1993) A spiking network model of short-term active memory. *Journal of Neuroscience* 13, 3406-3420