Theorie neuronaler Assoziativspeicher

Lokales Lernen und iteratives Retrieval von Information

Inaugural-Dissertation

zur

Erlangung des Doktorgrades der

Mathematisch-Naturwissenschaftlichen Fakultät

der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Friedrich Tobias Sommer

aus

Stuttgart

1993

Inhaltsverzeichnis

1 Einführung 5
- 1.1 Übersicht 5
- 1.2 Assoziative Datenspeicherung 6
  - 1.2.1 Retrieval – Adressierung 6
  - 1.2.2 Adressierungsfehler und Retrievalfehler 7
  - 1.2.3 Autoassoziation 8

2 Neuronale Assoziativspeicher 9
- 2.1 Neuronale Netzwerke 9
- 2.2 Retrieval 10
  - 2.2.1 Das Neuronenmodell 10
  - 2.2.2 Binäre Aktivitätsmuster 10
  - 2.2.3 Netzwerktopologie 11
- 2.3 Speicherprozess 12
  - 2.3.1 Synaptische Lernregeln 12
  - 2.3.2 Speicherverfahren 13
- 2.4 Verteilte Speicherung 13
- 2.5 Gehirnmodelle 14

3 Informationstheorie 16
- 3.1 Zum Informationsbegriff 16
- 3.2 Binärer Informationskanal 17
- 3.3 Übertragung binärer Muster 19
- 3.4 Zufallsgenerierung von Mustermengen 19
- 3.5 Übertragung von Mustermengen 21
- 3.6 Speicherkapazität 21
  - 3.6.1 Retrievaltestschema 21
  - 3.6.2 Kapazitätsdefinitionen 22
  - 3.6.3 Hifi-Qualitätskriterium 23

4 Einschrittretrieval 25
- 4.1 Statistische Analyse 25
- 4.2 Signal-zu-Rausch-Verhältnis 26
- 4.3 Adressierung 26
  - 4.3.1 Festgehaltene Musteraktivierung 26
  - 4.3.2 Feste Anzahl der On-Neurone 27
- 4.4 Binäre Speicherung mit der Hebbregel 27
  - 4.4.1 Fehlerwahrscheinlichkeiten 27

4.4.2 Vernachlässigung synaptischer Korrelationen . . . 28
4.4.3 Vergleich unterschiedlicher Adressfehler . . . 29
4.4.4 Heteroassoziation . . . 29
4.4.5 Autoassoziation . . . 31
4.4.6 Literatur . . . 32
4.5 Inkrementelle Speicherung . . . 33
4.5.1 Berechnung des Signal-zu-Rausch-Verhältnis . . . 34
4.5.2 Signal-zu-Rausch-Verhältnis für spezielle Lernregeln . . . 35
4.5.3 Optimale Lernregel . . . 35
4.5.4 Hebbregel und Spinregel . . . 36
4.5.5 Signal-zu-Rausch-Verhältnis und Lernregeln . . . 37
4.5.6 Heteroassoziation . . . 38
4.5.7 Autoassoziation . . . 39
4.6 Diskussion . . . 40
4.6.1 Heteroassoziation . . . 40
4.6.2 Autoassoziation . . . 42

5 Fixpunktkapazität **43**
5.1 Zusammenfassung . . . 43
5.2 Fixpunkt-Ausleseverfahren . . . 43
5.3 Informationskanal . . . 44
5.4 Hifi-Qualitätskriterium . . . 45
5.4.1 Perfekte Adressierung . . . 45
5.4.2 Unscharfe Adressierung . . . 46
5.5 Auslesefehlerwahrscheinlichkeiten . . . 47
5.6 Binäre Speicherung . . . 48
5.7 Inkrementelle Speicherung . . . 49
5.8 Diskussion . . . 50

6 Iteratives Retrieval **51**
6.1 Fragestellung . . . 51
6.2 Schwellenstrategien . . . 51
6.3 Statistische Analyse des zweiten Schrittes . . . 52
6.3.1 Hifi-Korrekturkapazität . . . 55
6.3.2 Fehlerwahrscheinlichkeit im endliches System . . . 56
6.4 Eigenschaften im endlichen System . . . 57
6.4.1 Binäres versus inkrementelles Lernverfahren, n=1900 . . . 58
6.4.2 Schwellenstrategien für binäre Speicherung, n=1900 . . . 60
6.4.3 Optimale Modellvariante, n=1900 . . . 63
6.4.4 Theorie und Experiment am Optimum . . . 66

7 Diskussion und Zusammenfassung **71**
7.1 Thermodynamische Methode . . . 71
7.1.1 Energiefunktion . . . 71
7.1.2 Spinglas-Theorie . . . 73
7.1.3 Analyse der Dynamik . . . 74
7.1.4 Die Gardner Methode . . . 74
7.2 Autoassoziation im Assoziativspeicher . . . 75

7.2.1 Asymptotische Ergebnisse . . . 75
7.2.2 Endliche Systemgröße . . . 76
7.3 Konsequenzen für die Anwendung . . . 77
7.4 Konsequenzen für die Hirnforschung . . . 78
7.5 Resumé . . . 79

A Binäre Hebbregel **86**

B Inkrementelle Hebbregel **90**

C Fixpunkt-Ausleseverfahren **92**

D Liste oft verwendeter Symbole **94**

Zusammenfassung

Die vorliegende Arbeit untersucht die Eigenschaften von lokalen synaptischen Lernregeln in Netzwerkmodellen binärer Schwellwertneurone als Assoziativspeicher. Im Mittelpunkt stehen Modelle mit iterativem Retrieval.

Um unterschiedliche Modellvarianten vergleichen zu können, behandeln wir Speicher- und Ausleseprozess mit Methoden aus der Informationstheorie als Übertragungskanal. Die Kanalkapazität pro Synapse – die Speicherkapazität – wird als Gütemaß verwendet: sie gibt abrufbare Information pro belegte Größe des Speichermediums (Anzahl der Synapsen) an.

Zunächst wird Analyse und Optimierung der synaptischen Lernregel bei inkrementeller Speicherung im Einschrittretrieval dargestellt. Die Speicherung spärlicher Muster mit der Hebbregel erreicht hierbei die höchste asymptotische Speicherkapazität von $K = 0.18$ Bit/Synapse. Die geklippte (binäre) Variante des inkrementellen Speicherprozesses arbeitet fast unverändert gut $K = 0.17$ Bit/Synapse..

Für diese bei Einschrittretrieval optimalen Modellvarianten wird ein neues Analyseverfahren entwickelt, welches mit elementaren kombinatorischen und statistischen Methoden ein Ausleseverfahren untersucht, bei dem ein großer Raum von Eingangsmustern nach Fixpunkten der Retrievaldynamik abgesucht und die so erhaltene Menge von Mustern als Ausleseergebnis interpretiert wird. Die asymptotische Speicherkapazität dieses Verfahrens – die Fixpunktkapazität – ist eine obere Schranke der Kapazität für iteratives Retrieval, daß bis zum Fixpunkt ausgeführt wird. Diese Schranken liegen genau doppelt so hoch wie die Kapazität des Einschrittretrieval.

Für das binäre Hebbmodell wird eine Variante des iterativen Retrieval eingeführt, bei der Ein- und Ausgangsmuster ab dem zweiten Schritt logisch geUNDet werden. Diese Variante weist für Anwendungen die günstigsten Eigneschaften auf: hohe Speicherkapazität, kurze Iterativszeiten und kein Auftreten der bei Mustervervollständigung störenden Zweierzyklen. Ihre asymptotische Analyse für Autoassoziation liefert jedoch Kapazitätswerte, die genau mit denen von Einschrittretrieval übereinstimmen. Asymptotisch bringt also die Iteration des Retrieval keine Verbesserung.

Simulationsexperimente werden bis zu Systemgrößen von $n = 20000$ durchgeführt. Während die Experimente mit inkrementelle Speicherung nur den halben Wert der asymptotischen Kapazität erreichen, liegt die Kapazität mit binärer Speicherung für Einschrittretrieval bereits bei leicht simulierbaren Systemgrößen auf 90% des asymptotischen Wertes. Mit iterativem Retrieval liegen die Kapazitätswerte, zunächst überraschend, sogar über dem asymptotischen Wert.

Für binäre Speicherung im endlichen System werden die ersten zwei Schritte von iterativem Retrieval analysiert. Die Theorie des zweiten Schrittes bestätigt unsere experimentellen Ergebnisse bei iterativem Retrieval: die Kapazität erreicht ein Maximum oberhalb des asymptotischen Wertes bei endlicher Systemgröße und fällt für weiter wachsende Systemgröße auf den asymptotischen Wert zurück. Im endlichen System erzielt also iteratives Retrieval durchaus eine Verbesserung gegenüber Einschrittretrieval.

Insgesamt weisen unsere Untersuchungen die UNDungs- Variante des iterativen Retrieval im binären Hebbmodell als die für Mustervervollständigung leistungsfähigste neuronale Realisierung des Assoziativspeichers aus. Im endlichen System ist damit eine Speicherkapazität von 0.19 Bit/Synapse in weniger als fünf Retrievalschritten möglich – gegenüber Einschrittretrieval bedeutet das eine Erhöhung um ca. 20%. Zudem kann dieser Maximalwert in Anwendungen stärker ausgeschöpft werden, da er bei Retrievalfehlern angenommen wird, die nur ca. 20% der Fehler am Optimum des Einschrittretrieval betragen.

Kapitel 1

Einführung

1.1 Übersicht

Das Charakteristikum eines Assoziativspeichers besteht darin, fehlertolerante Zuordnungen zwischen Datenpaaren (hier Paaren binärer Muster) zu leisten. Die folgenden Abschnitte dieses Kapitels nehmen eine Klassifizierung solcher Zuordnungsaufgaben in Hetero- und Autoassoziation vor und definieren in diesem Zusammenhang den Begriff der Fehlertoleranz.

Eine kurze Einführung in einfache neuronale Assoziativspeichermodelle mit lokalen Lernregeln gibt das zweite Kapitel. Die Speicherung findet durch Änderung der synaptischen Verbindungen statt, welche Eingangssignale gewichtet zu den Neuronen übertragen – die Synapsen stellen also in diesen Modellen das Speichermedium dar. Ausgelöst werden die synaptischen Modifikationen im Speicherprozess durch Anlegen der Solldaten. Die synaptische Lernregel ist für die Speicherung ein wichtiges Modellmerkmal: sie bestimmt, welche Änderung der Synapse durch die anliegenden Sollwerte ausgelöst wird. Im Ausleseprozess oder Retrieval[1] arbeiten die Modellneuronen als einfache binäre Schwellwertschalter. Die Signale, welche sie über die Synapsen erreichen, werden aufaddiert und zu einem neuronalen Ausgangssignal ausgewertet. Abhängig davon, ob bei einer Speicheranfrage jedes Neuron nur einmal oder mehrere Male hintereinander eine Auswertung vornimmt, unterscheiden wir Einschrittretrieval und iteratives Retrieval. Die Untersuchung von iterativem Retrieval bildet den zentralen Gegenstand der vorliegenden Arbeit.

Um verschiedene Modellvarianten bei einer Aufgabe wie z. Bsp. Mustervervollständigung vergleichen zu können, ist ein Gütemaß notwendig, das gespeicherte Datenmenge und im Retrieval auftretende Fehler bezüglich der Größe des benötigten Speichermediums berücksichtigt. Hierzu werden im dritten Kapitel Einspeicher- und Retrievalvorgang mit elementaren Methoden der Informationstheorie als Informationkanal behandelt und die Speicherkapazität eingeführt. Die Speicherkapazität mißt die abrufbare Information pro Synapse in Einheiten Bit/Synapse, bei Mustervervollständigung wird die Information in den Adressmustern subtrahiert. Dieses Gütemaß erlaubt den Vergleich von Modellen, die sich hinsichtlich der Parameter Systemgröße, Anzahl und mittlere Aktivität in den Lernmustern, mittlere Verrauschung der Adressmuster, etc. unterscheiden können.

Das vierte Kapitel referiert zunächst die Signal-zu-Rausch-Analyse des Einschrittretrieval im Assoziativspeicher. Diese Analyse weist die Hebbsche Lernregel in der Kombination mit spärlichen Lernmustern als Modell mit besonders hoher Speicherkapazität aus. Varianten

---

[1] Wir behalten hier das englische Wort bei. Es umfaßt sinngemäß die Begriffe ‘Auslesen’ und ‘Wiederauffinden’ eines Eintrags, ist aber durch eines dieser Worte allein nicht hinreichend beschrieben.

dieses Modells mit iterativem Retrieval für Mustervervollständigung stehen im Zentrum der weiteren Teile dieser Arbeit.

Im fünften Kapitel kann sich der Leser beim ersten Durchgehen der Arbeit auf die vorangestellte Zusammenfassung beschränken. Dieses Kapitel stellt insofern einen Einschub dar, als daß dort ein Ausleseverfahren und kein Retrievalverfahren analysiert wird. Das Ausleseverfahren extrahiert durch eine große Anzahl von Speicheranfragen die Menge der Muster, die durch Hebbsche Speicherung dem System als Fixpunkte der Retrievaldynamik eingeprägt wurden. Es eignet sich daher nicht zum Wiederauffinden eines einzelnen Musters. Seine Speicherkapazität – die Fixpunktkapazität – stellt aber eine obere Schranke für die Kapazität des iterativen Retrieval (das bis zum Fixpunkt geführt wird) dar.

Die Ergebnisse des fünften Kapitels weisen darauf hin, daß durch iteratives Retrieval höhere Kapazitätswerte möglich sind. Im sechsten Kapitel findet die Untersuchung des iterativen Retrieval statt. Es werden Strategien zur Schwellenregelung und zum Iterationsabbruch eingeführt, die sich im Experiment bewährt haben. Für die beste dieser Strategien ist die asymptotische Analyse sowie die Analyse des zweiten Iterationsschrittes im System endlicher Größe durchgeführt. Es folgt eine ausführliche Darstellung der Eigenschaften von iterativem Retrieval anhand zahlreicher Diagramme der analytischen und experimentellen Ergebnisse. Unsere Untersuchungen geben Antwort auf die Frage, in welchen Fällen sich iteratives Retrieval gegenüber Einschrittretrieval im Hebbmodell lohnt.

Schließlich ordnet das siebte Kapitel unsere Ergebnisse in die Resultate aus der Literatur ein. Zum Verständnis jener Resultate werden die wichtigsten Analysemethoden rückgekoppelter Assoziativspeichermodelle aus der statistischen Physik knapp umrissen. Wir diskutieren die Konsequenzen der vorliegenden Untersuchungen, einerseits in Bezug auf Anwendungen und andererseits im Hinblick auf Hirntheorie.

Der Themenkreis assoziativer Speicherung in neuronalen Netzen ist durch unterschiedliche Anforderungen der Anwendungen und durch die Vielzahl von Modellen nur sehr schwer kohärent und umfassend darstellbar. Diese Arbeit verfolgt den Pfad zu möglichst effizienter Mustervervollständigung, der auf Hebbspeicherung und iteratives Retrieval führt und behandelt dabei zwangsläufig andere Aspekte weniger ausführlich. In einem kurzen Resumé am Schluß sind Fragestellungen und Erkenntnisse nocheinmal so aufgereiht, wie sie den Anstoß zu dieser Arbeit gaben und ihren Verlauf bestimmt haben.

Es wurde versucht, die Notation in dieser Arbeit durchgehend konsistent zu halten. Die wichtigsten der verwendeten Symbole sind im Anhang D in einer Liste mit kurzen Erklärungen zusammengestellt.

1.2 Assoziative Datenspeicherung

Jede Form von Datenspeicherung setzt ein physikalisches System als *Speichermedium* mit entsprechenden Vorschriften zur Kodierung der Daten voraus, die den Schreib- und Lesevorgang spezifizieren. Die charakteristische Eigenschaft des Speichermediums besteht dann darin, Zustandsänderungen, die ihm im Schreibvorgang eingeprägt wurden, stationär beizubehalten, so daß sie zu einem späteren Zeitpunkt dem Lesevorgang zugänglich sind.

1.2.1 Retrieval – Adressierung

Um kleine Inhaltssegmente in gespeicherten Datenmengen abrufen zu können, muß der Speicherinhalt segmentiert werden und die einzelnen Segmente mit Adressen versehen werden.

Mit *Retrieval* bezeichnen wir den selektiven Lesevorgang, der durch Eingabe einer Segmentsadresse auf ein einzelnes Inhaltssegment zugreift.

Im weiteren sind Inhaltssegment sowie Segmentadresse Ketten zweiwertiger Variablen $x \in \{a, 1\}^n$ mit $a \in [-1, 0]$, also binäre Daten oder *binäre Muster.* Ein solches Muster wird später im neuronalen Assoziativspeicher repräsentiert durch Neuronenzustände. Wir bezeichnen oft die Menge $\{x_i : x_i = 1\}$ als die *On-Komponenten* eines neuronalen Aktivitätsmusters und entsprechend die Menge $\{x_i : x_i = a\}$ als *Off-Komponenten*, vgl. Abschnitt 2.2.1.

Die Aufgabe eines Datenspeichers läßt sich spezifizieren durch Mengen vorgegebener *Adressmuster* $\mathcal{S}^A$ und *Inhaltsmuster* $\mathcal{S}^C$, die zu Assoziationspaaren gruppiert sind:

$$\{(x^1, y^1), ..., (x^M, y^M) : x^k \in \mathcal{S}^A, y^k \in \mathcal{S}^C\}. \tag{1.1}$$

Man unterscheidet grundsätzlich zwei verschiedene Organisationen in Datenspeichern: In einem *Listenspeicher* werden die Adress- und Inhaltsmuster während des Speicherprozess direkt in den Zustandsraum des Speichermediums abgebildet und ihm aufgeprägt. Bei Datenzugriff wird der Adressenteil des listenartig organisierten Speichermediums nach der passenden Adresse abgesucht und das gewünschte Inhaltsmuster ausgelesen (wie z. Bsp. beim Telefonbuch). In einem *Abbildungsspeicher* wird bei jedem Datenzugriff die Adress-Inhaltsmuster Abbildung errechnet. Hierzu muß im Speicherprozess der notwendige Algorithmus im Speichermedium abgelegt worden sein, mit dessen Hilfe das Adressmuster auf das Inhaltsmuster transformiert werden kann. Das Problem bei diesem Speichertyp besteht gewöhnlich darin, für einen einzuspeichernden Datensatz einen Algorithmus zur Berechnung der Abbildung zu finden. Bei einem Speicher neuronaler Architektur handelt es sich um einen Abbildungsspeicher.

1.2.2 Adressierungsfehler und Retrievalfehler

Fehlertoleranz bei *assoziativer Speicherung* setzt zwei Eigenschaften der Adressierung voraus: erstens muß auf dem (diskreten) Adressraum ein Abstandsmaß definiert sein und zweitens muß Redundanz in der Adressierung vorliegen, d.h. nicht alle im Adressraum liegenden Adressen dürfen unterschiedlichen Datensegmenten zugewiesen werden. Mit diesen Voraussetzungen kann ein Datensegment auch durch ein Eingangsmuster abgerufen werden, welches nicht exakt einer eingespeicherten Adresse identisch ist, sondern lediglich einen geringen Abstand zu ihr besitzt.

Abstände zwischen Mustern
Ein natürliches Abstandsmaß bei binären Mustern bildet die *Hammingdistanz* $h(x, \hat{x})$. Sie ist definiert als die Anzahl voneinander abweichender Komponenten zwischen x und $\hat{x}$. Die Abweichung zwischen $\hat{x}$ von x kann noch genauer beschrieben werden durch Fehlerzahlen der beiden möglichen *Fehlerttypen. On-Fehler* werden in der Indexgruppe $\{i : x_i = 1\}$, den On-Komponenten von x gezählt, entsprechend *Off-Fehler* in den Off-Komponenten mit $\{i : x_i = a\}$. Im weiteren verwenden wir oft *Fehlerverhältnisse*: $\tilde{\beta}$ ist die Anzahl On-Fehler pro Anzahl der On-Komponenten und $\tilde{\alpha}$ die Anzahl Off-Fehler pro Anzahl der Off-Komponenten.

Fehlertoleranz in der Adressierung
Über ein solches Abstandsmaß kann eine große Menge von Eingangsmuster ein gewünschtes Inhaltsmuster eindeutig adressieren[2]. Ein Speicher arbeitet *fehlertolerant* , wenn er Adressie-

---

[2] Beispielsweise adressieren alle Muster $\hat{x}$ mit der Eigenschaft $h(\hat{x}, x^k) < h(\hat{x}, x^l)$ für alle $x^l \neq x^k \in \mathcal{S}^A$ das Muster x^k eindeutig.

rungsfehler zuläßt, indem ein Eingangsmuster $\hat{x} \notin \mathcal{S}^A$ auf das Inhaltsmuster abbildet wird, dessen zugehöriges Adressmuster benachbart ist.

Eine solche Toleranz gegenüber unscharfen Eingaben ist dem psychologischen Begriff des Assoziierens verwandt. Für den Assoziativspeicher hat diese Verwandschaft hat nicht nur zur Namensgebung geführt, sondern auch die ersten technischen Realisierungen [St61] beeinflußt.

Retrievalfehler
Assoziativspeichermodelle, wie sie im weiteren betrachtet werden, verhalten sich bezüglich der Eingabe fehlertolerant. Es können aber auch Retrievalfehler auftreten, also Abweichungen zwischen Ausgangsmustern und Inhaltsmustern. Bei nicht fehlerfrei arbeitendem Speicher kann über die Fehlerverhältnisse im Ausgangsmuster die Güte des Retrieval beurteilt werden. Wie man intuitiv erwarten würde, steigen mit größerwerdenden Adressierungsfehlern auch die Retrievalfehler.

1.2.3 Autoassoziation

Die Zuordnungsaufgabe zwischen beliebigen Adressmustern und Inhaltsmustern (1.1), nennt man *Heteroassoziation.* Den Spezialfall, daß Adress- und Inhaltsmuster identisch sind, bezeichnet man als *Autoassoziation* oder auch als Inhaltsadressierung. Falls in der Adressierung Fehlertoleranz gewährleistet ist, kann *Mustervervollständigung* stattfinden, der Speicher vervollständigt dann ein fehlerhaftes Muster $\hat{x}^k$, daß er als Eingangsmuster erhält, zum gespeicherten Inhaltsmuster x^k.

Kapitel 2

Neuronale Assoziativspeicher

2.1 Neuronale Netzwerke

Es gibt mittlerweile eine große Zahl von guten einführenden Büchern über neuronale Netzmodelle, wie [Ko77, At89, DHS91, HKP91]. Wir werden daher in diesem Abschnitt eine nur sehr knappe allgemeine Einführung neuronaler Netzmodelle geben und uns dann ganz auf assoziative Speicher beschränken.

Neuronale Netzwerke sind signalverarbeitende Systeme, die durch drei Komponenten charakterisiert werden können: Neuronenmodell, Synapsenmodell und vorgegebene Netztopologie der Synapsen.

Ein neuronales Netzwerkmodell besteht aus einer großen Anzahl gleichartiger sowohl untereinander als auch mit Ein- und Ausgangsschnittstellen hochvernetzter Prozessoren. Die Prozessoren, nach ihrem biologischen Vorbild *Neurone* genannt, berechnen aus Eingangssignalen ihre Aktivitätszustände, die das Ausgangsmuster des Netzwerks bilden. Die Neurone besitzen im einfachsten Fall[1] keine zusätzlichen inneren Zustände. Ein Neuron erhält Signale vom Eingang des Netzwerks und, je nach Netztopologie von anderen Neuronen, über in der Übertragungsstärke modifizierbare Verbindungen, die *Synapsen.* Im Rahmen einer vorgegebenen Netztopologie formt und organisiert sich die globale synaptische Vernetzungsstruktur, in der die Neurone eingebunden sind, durch Rechenprozesse in den Synapsen, die auch durch das Netz propagierte Signale auswerten.

Es finden also zwei unterschiedliche Rechenoperationen in den Neuronen und in den Synapsen statt. Diese Operationen finden in den meisten Modellen in zeitlich getrennten Phasen statt: In der Lernphase werden die zu lernenden Daten als Sollwerte an Ein- und Ausgangsleitungen angelegt. Jede Synapse stellt ihre Verbindungsstärke, abhängig von anliegenden Signalen gemäß einer Lernregel[2] ein. Das in Abschnitt 1.2.1 angesprochene Problem des Abbildungsspeichers, einen geeigneten Algorithmus zu finden, wird also in neuronalen Speichermodellen im Lernvorgang gelöst.

Nach abgeschlossener Lernphase kann das Retrieval durch den Rechenvorgang der Neurone stattfinden. An den Eingangsfasern anliegende Signale werden über die Synapsen gewichtet an die Neurone weitergeleitet. Jedes Neuron verrechnet seine Eingangssignale zu einem Ausgangssignal, welches je nach Netztopologie entweder direkt auf eine Ausgangsfaser geschaltet ist, oder über Synapsen zu weiteren Neuronen weitergeleitet wird.

---

[1] In Modellen ohne Refraktärphase.

[2] In vielen Synapsenmodellen, wie auch in den von uns verwendeten, sind diese Lernregeln lokal. Sie werten also lediglich die lokal an der Synapse verfügbaren Signale aus. Im Gegensatz dazu gibt es auch nichtlokale Lernregeln, zum Beispiel die Pseudoinverse-Lernregel.

2.2 Retrieval

Beim Datenzugriff wird ein Adressmuster an die Eingangsfasern angelegt und nach vorgegebenem Schema der Rechenvorgang, der *Updateprozess* der Neurone zugelassen.

2.2.1 Das Neuronenmodell

Im Update Prozess wertet ein Neuron j seine momentan anliegenden Eingangssignale aus. Es bildet die über die Synapsen c_{ij} gewichtete Summe aller einlaufenden Aktivitäten[3], das *dendritische Potential*

$$d_j := \sum_i c_{ij} x_i \tag{2.1}$$

und aktualisiert seinen *Aktivitätszustand* $\tilde{y}_j$ gemäß der *Update-Gleichung*

$$\tilde{y}_j = f(d_j - \Theta_j). \tag{2.2}$$

Die Funktion $f(x)$ heißt *neuronale Transferfunktion* und ist in fast allen Modellen monoton steigend. Θ_j ist ein von außen einstellbarer Offset, der hier für alle Neurone gleich eingestellt wird[4].

Wir beschränken uns auf Neuronenmodelle mit sprungförmiger Transferfunktion der Form: $f(x) = 1$ für $x \geq 0, f(x) = a$ sonst mit $a < 1$. Sie heißen *binäre Neurone*, weil ein Neuron genau zwei unterschiedliche Aktivitätszustände annehmen kann. Analog zur biologischen Terminologie nennen wir den Aktivitätszustand $\tilde{y}_j = 1$ *aktiv* oder *feuern* und entsprechend $\tilde{y}_j = a$ *passiv.* Der Parameter Θ wird oft als Feuerschwelle oder kurz als *Schwelle* bezeichnet. Zu diesen Bezeichnungen paßt natürlich besonders gut die Parameterwahl $a = 0$, wir nennen daher dieses Modell das *biologische Modell.*

2.2.2 Binäre Aktivitätsmuster

Sind auch als Eingangssignale nur Aktivitätswerte aus der Menge $\{a, 1\}$ zugelassen, so können die Eingangssignale, sowie die Aktivität aller Neurone im Netz zu einem m- bzw. n-dimensionalen binären Muster zusammenfaßt werden. Die *Aktivierung* $p(x)$ *eines Musters* x ist definiert als das Verhältnis zwischen der Anzahl aktiver Neurone, der On-Neurone $k(x) := \sharp\{i : x_i = 1\}$ und ihrer Gesamtzahl

$$p(x) := k(x)/n. \tag{2.3}$$

Die Aktivierung entspricht also der Wahrscheinlichkeit, mit dem eine Musterkomponente von x aktiv ist. Die *Aktivität* eines Musters ist definiert als der Mittelwert einer Komponente des Musters x

$$\mu(x) := \frac{1}{n} \sum_i x_i = p(x) + [1 - p(x)]a. \tag{2.4}$$

Die Aktivität ist im Gegensatz zur Aktivierung vom Zahlenwert des Parameters a abhängig, für $a = 0$ stimmen beide Definitionen überein.

---

[3] Das Synapsengewicht c_{ij} verbindet Eingang i der Eingangsschnittstelle mit Neuron j.

[4] Es wurden Modelle vorgeschlagen, in denen Θ_j für jedes Neuron individuell eingestellt wird, z. Bsp. in [WD91].

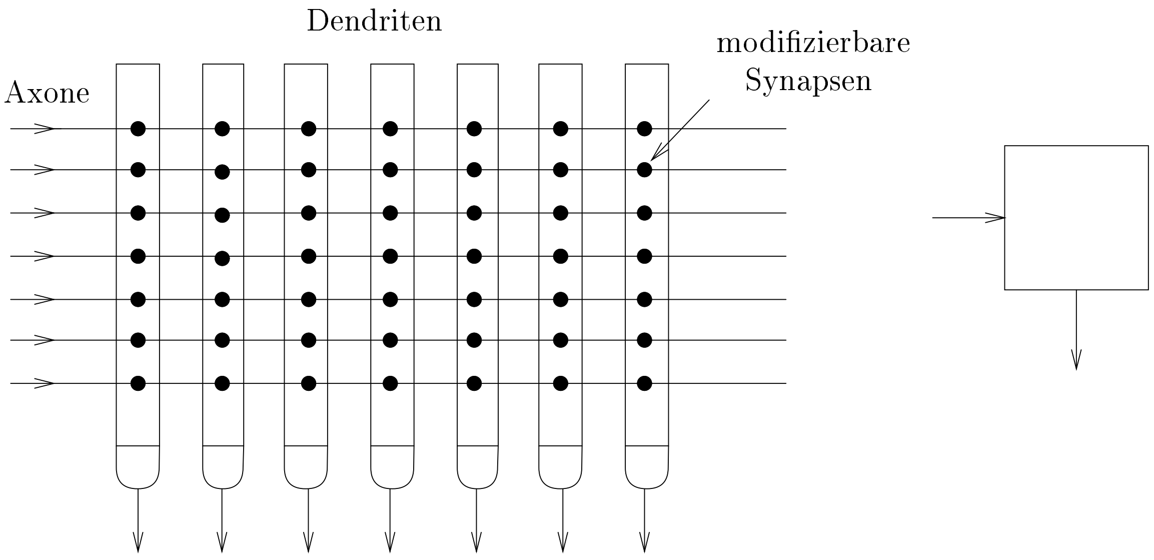

Abbildung 2.1: Schema eines neuronalen Assoziativspeichers

2.2.3 Netzwerktopologie

Bei der allereinfachsten Netzwerkvariante werden die Eingangsfasern über die Synapsenmatrix an die Neurone geleitet, das Aktivitätsmuster der Neurone bildet dann direkt den Ausgangsvektor. Diese Netzwerkmodelle und Kaskaden aus mehreren solchen *Neuronenschichten* bilden die Klasse der *feed-forward Modelle.*

Die rückgekoppelte Variante unterscheidet sich dadurch, daß die Anzahl von Eingangs- und Ausgangsfasern gleich ist und beim Retrieval jede Ausgangsfaser mit einer Eingangsfaser gekoppelt werden kann, vgl. Abbildung 2.2.

Da beide Varianten sich nur im Retrievalprozess unterscheiden, kennzeichnen wir sie mit *Einschrittretrieval* beziehungsweise *iterativem Retrieval.*

Einschrittretrieval

Das Ausgangsmuster von Einschrittretrieval wird durch die Ausgangsaktivitäten $\tilde{y}_j$ der Neurone gebildet, nachdem jedes Neuron für das anliegende Eingangsmuster einen Update ausgeführt hat. Die Schwelle wird für alle Neurone auf den selben Wert Θ gesetzt, dem einzigen Parameter, der beim Retrieval von außen (vom Benutzer) einzustellen ist.

Iteratives Retrieval

Bei Rückkopplung des Ausgangsmusters auf den Eingang sind zwei Varianten zu unterscheiden, je nachdem wie das Update verschiedener Neurone im Netz koordiniert wird.

Bei *sequentiellem Update* wird pro Iterationsschritt nur ein zufällig mit Gleichverteilung ausgewähltes Neuron zum Update zugelassen, alle anderen Neurone behalten ihre bisherigen Ausgangswerte bei.

Paralleles Update bedeutet vollständiges Einschrittretrieval in jedem Iterationsschritt, wir werden im weiteren diese Variante des iterativen Retrieval betrachten. Hier ist für jeden Iterationsschritt der Schwellenwert von außen einzustellen und zusätzlich durch eine Abbruchbedingung der Ausstieg aus der Iterationsschleife festzulegen. Eine solche Vorschrift nennen wir *Schwellenstrategie.* Die in dieser Arbeit vorgeschlagenen und untersuchten Strategien werden in Abschnitt 6.2 beschrieben.

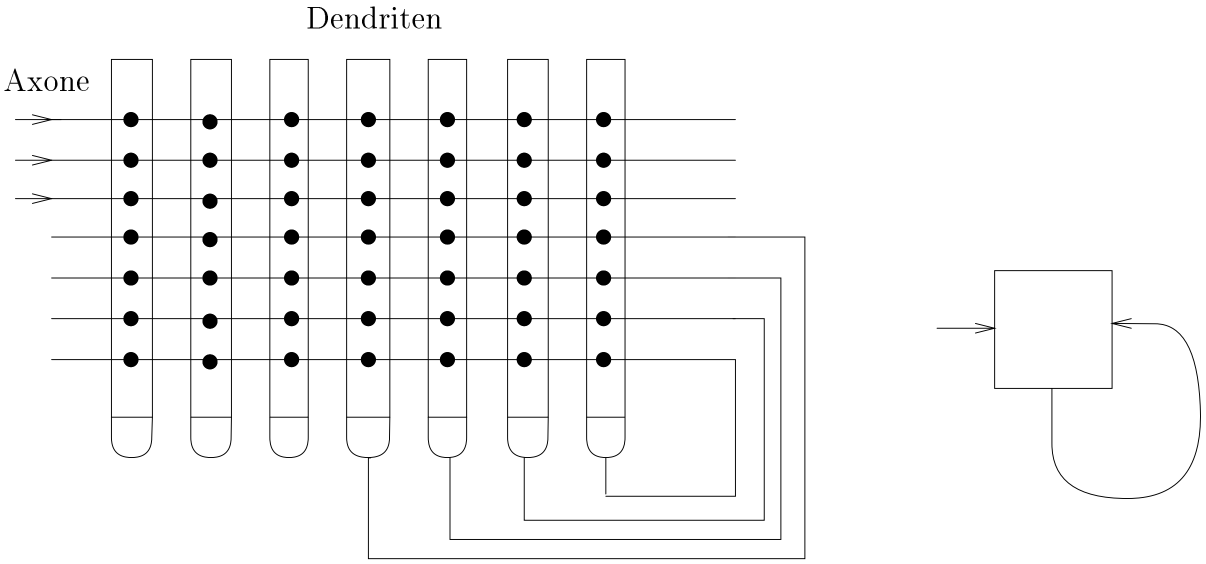

Abbildung 2.2: Rückgekoppelter Assoziativspeicher

2.3 Speicherprozess

Im Speicherprozess oder Lernprozess wird die synaptische Verbindungsmatrix, das Speichermedium geformt. Nacheinander wird jedes zu lernende Musterpaar (x^l, y^l) an Ein- und Ausgängen des Speichers angelegt, so daß jede Synapse c_{ij} pre- und postsynaptisch die Sollwerte der Aktivität (x_i^l, y_j^l) erhält, die bei idealem Retrieval dieses Musterpaares an Ein- und Ausgang anliegen sollte. Die Neurone sind während des Speicherprozesses entweder in Rückwärtsrichtung leitend durchgeschaltet, oder die Ausgangssollwerte werden über getrennte Lernfasern an die Synapsen geleitet.

2.3.1 Synaptische Lernregeln

Für ein Paar solcher Aktivitäts-Sollwerte (x, y) bestimmt eine *lokale synaptische Lernregel* $R(x, y)$ den jeweiligen Betrag der synaptischen Änderung. Bei binären Mustern gibt es vier verschiedene Konstellationen der pre- und postsynaptischen Sollwerte: $(a, a), (a, 1), (1, a)$ und $(1, 1)$. Eine lokale synaptische Lernregel ist also vollständig durch vier Zahlen bestimmt:

$$R = (r_1, r_2, r_3, r_4) \tag{2.5}$$

Zwei Lernregel, die in den letzten Jahren besondere Aufmerksamkeit auf sich gezogen haben sind:

- Die *Hebbregel*: $H := (0, 0, 0, 1)$, sie erhöht den Synapsenwert bei gleichzeitigem pre- and postsynaptischem Feuern. In seinem berühmten neurophysiologischen Postulat schlug D. O. Hebb [He49] diese synaptische Änderungsvorschrift im Nervensystem vor.

- Die Hopfield oder *Spinregel*: $S := (1, -1, -1, 1)$. Hier wird der Synapsenwert in allen Fällen übereinstimmender pre- und postsynaptischer Aktivität erhöht. Diese Regel wurde auch im ursprünglichen Hopfieldmodell [Hd82] benutzt.

Beide Regeln sind Produktregeln, d. h. von der Form: $R(x,y) = xy$. Im biologischen Modell mit $a = 0$ erhalten wir die Hebbregel, im Isingmodell mit $a = -1$, die Hopfieldregel und oft, beispielsweise in [HSKH88], werden beide als Hebbsches Lernen bezeichnet. Da aber Hebb in der Orginalformulierung eindeutig gleichzeitiges Feuern als Auslöser für die Erhöhung eines Synapsenwertes angibt, behalten wir die Unterscheidung bei. Der Psychologe Hebb war von den damals neuen elektrophysiologischen Experimenten inspiriert, welche den unterschiedlichen Charakter der beiden Aktivitätszustände bei den Nervenzellen zeigten. Auch sein Ansatz, psychologische Phänomene wie Gedächtnisinhalte durch interagierende Gruppen aktiver Neurone – den sogenannten 'Cell assemblies' – zu erklären[5], ist nur im $a = 0$ Bild denkbar. Es ist wenig verwunderlich, daß in der biologisch orientierten Literatur das (a=0)-Modell dominiert. Hingegen ist es wegen der Symmetrie des Isingmodells bezüglich der Mustertranformation $1 \to -1, -1 \to 1$ typisch für die physikalisch orientierte Literatur, die Aktivitätszustände eher als äquivalent zu betrachten.

2.3.2 Speicherverfahren

Wir untersuchen hier auschließlich Lernverfahren mit einem Lernschritt, die also nach einmaligem Anlegen eines jeden Musters beendet sind[6]. Zwei unterschiedliche Verfahren werden im weiteren behandelt:

- Das *inkrementelle Lernverfahren*, bei dem die synaptische Matrix gegeben ist durch

$$c = (c_{ij}) := \sum_{k=1}^{M} R(x_i^k, y_j^k). \tag{2.6}$$

- Das *binäre Lernverfahren*, bei dem die synaptische Matrix $\bar{c}$ aus c gewonnen wird durch

$$\bar{c}_{ij} := \operatorname{sgn}(c_{ij})\,, \tag{2.7}$$

  mit $\operatorname{sgn}(0) := 0$.

Für die Hebbregel erhalten wir die synaptischen Matrizen durch die Gleichungen

$$\begin{aligned} c_{ij} &= \sum_{k=1}^{M} x_i^k y_j^k && (2.8) \\ \bar{c}_{ij} &= 1 - \prod_{k=1}^{M} (1 - x_i^k y_j^k) = \min(1, c_{ij}) && (2.9) \end{aligned}$$

2.4 Verteilte Speicherung

Einer der Gründe, weshalb Systeme mit neuronaler Architektur nach einer Dekade geringer Beachtung in den siebziger Jahren zu neuem Leben erweckt wurden, ist die Tatsache, daß sich

---

[5] Mit dieser Theorie trat er dem damals in der Psychologie verbreiteten Behaveriorismus entgegen. Eine wichtige Anregung für Hebb waren Arbeiten von R. Lorente de Nó, der Rückkopplungsbahnen im Gehirn fand und die damals herrschende Meinung widerlegte, daß das Gehirn gleich wie der Verdauungsapparat unidirektional arbeitet.

[6] Im Kontrast dazu steht zum Beispiel beim Perceptron der iterative Lernvorgang mit einer Dreiterm-Lernregel. Diese Lernregel verändert die Synapse abhängig von Eingangs-, Soll- und Istwert. Bei mehrfachem Durchlaufen der Lernmuster im iterativen Speicherprozess kann dann die Übereinstimmung zwischen Soll- und Istwert sukzessive verbessert werden.

in den Computerwissenschaften 'distributed processing' als unverzichtbares neues Paradigma herauskristallisierte. Inwieweit ist nun schon an den einfachen Modellen, die in diesem Kapitel eingeführt wurden, die Eigenschaft von verteilter Speicherung zu beobachten ?

Bei Heteroassoziation werden durch lokale Lernregeln Korrelationen zweiter Ordnung zwischen Adressmuster und Inhaltsmuster gespeichert. Bei der Hebbregel inkrementiert jedes Paar aktivierter Neurone (x_i^k, y_j^k) die Synapse c_{ij}. Bei Autoassoziation werden Autokorrelationen zweiter Ordnung der Musteraktivität gespeichert. Bei der Hebbregel verursacht jedes Paar aktivierter Neurone im Inhaltsmuster die Änderung einer Synapse. Man könnte also bereits von verteilter Speicherung sprechen, sobald das Lernen eines Musterpaares nichtlokale Veränderungen in der Speichermatrix hervorruft, also bei der Hebbregel, sobald mindestens ein beteiligtes Muster mehr als ein aktives Neuron enthält. Hierbei handelt es sich jedoch um eine reine Duplizierung und deshalb wird der Begriff 'verteilte Speicherung ' hier noch strenger gefaßt: Es wird gefordert, daß fast jedes Element der Speichermatrix Information über viele der gespeicherten Musterpaare enthält. Für beliebige lokale Lernregeln erhält man verteilte Speicherung, wenn die Muster in den Mengen der Adress- und Inhaltsmuster überlappen, d.h. gemeinsame On-Neurone besitzen. Dann nämlich setzt sich die Summe $\sum_k R(x_i^k, y_j^k)$ für die meisten Indexpaare i, j aus mehreren nichtverschwindenden Summanden zusammen. Wie in der Holographie wird ein Inhaltssegment (ein Musterpaar) verteilt auf das Speichermedium geschrieben und verschiedene Inhaltssegmente sind überlagert.

2.5 Gehirnmodelle

Die biologische Terminologie dieses Kapitels fordert sofort zur Frage heraus, was diese technisch anmutenden Speichermodelle mit dem biologischen Nervensystem gemeinsam haben. Daher wollen wir hier, wenn auch wenig erschöpfend, auf die Entwicklung in der Modellbildung des biologischen Nervensystems eingehen. Am Ende der Arbeit wird dieser Aspekt in Abschnitt 7.4 nochmals aufgegriffen.

Durch Färbetechniken, die nur einen kleinen Bruchteil der auf einem Mikroskopschnitt eng verknäulten und sonst beinahe durchsichtigen Gehirnzellen sichtbar machen, kennt man seit der Jahrhundertwende schon sehr genau die Anatomie von Nervenzellen. Der Zellkörper besitzt zwei Fortsätze, den stark verzweigten Dendritenbaum und die eher lineare wenig verzweigte Axonfaser.

Bald darauf wurde es technologisch möglich, die elektrischen Vorgänge auf den Zellmembranen zu messen. Ein biologisches Neuron befindet sich grob in einem von zwei möglichen Aktivitätszuständen. Bei kleinen kleinen Änderungen der Membranspannung bleibt die Zelle passiv, durch eine elektrische Reizung am Dendriten kann ein Feuern des Neurons ausgelöst werden; es wird dann vom Zellkörper ein elektrischen Impuls erzeugt, ein sogenannter Spike, der entlang der Axonfaser wandert. Beim Zusammenspiel mehrerer Neurone durch synaptische Kontakte zwischen Axonen und Dendriten spielt die zeitlich kurz begrenzte Signalform des Spike eine wichtige Rolle: die endlichen Laufzeiten auf der Axonfaser verursachen definierte Verzögerungen bei der Signalübertragung.

Meistens treten an einem Neuron jedoch keine Einzelspikes auf, sondern längere Ruheperioden welchseln sich mit Folgen von kurz aufeinanderfolgenden Spikes ab – sogenannten 'spike trains'.

Es wurden mathematische Modelle der Zellmembran entwickelt [HH52], die durch einen Satz von Differentialgleichungen sehr detailliert die elektrischen Vorgänge auf der Zellwand beschreiben. Wegen der hohen Komplexität dieser Modelle sind jedoch schon Netze aus mehreren Neuronen kaum noch handhabbar. Um kollektives Verhalten von Neuronen überhaupt beschreiben zu können, ist ein hohes Maß an Abstraktion, an Vereinfachung der biologischen Vorgänge notwendig: es werden größtenteils Modelle verwendet, die am Neuron anstatt einzelner Spikes nur eine mittlere Spikerate beschreiben – 'spike rate coding models'. In der einfachsten biologischen Modellierung wird die An- oder Abwesenheit eines 'spike train' als binäre Variable beschreiben. Solche Modelle entspechen gerade den hier beschriebenen Netzen binärer Schwellenneurone.

Da die Raten der Spikes bei vielen Neuronenarten kontinuierlich variieren können, gibt es detailliertere Modelle, welche die mittlere Spikerate mit einem kontinuierlichen Aktivitätsparameter beschreiben. Trotzdem gab und gibt es für einige Gebiete im Gehirn (Kleinhirn, Cortexmodelle) Modellvorschläge mit Assoziativspeichern unter Verwendung binärer Schwellenneuronen [Ma91].

Kapitel 3

Informationstheorie

Bisher wurden verschiedene Aufgaben assoziativer Datenspeicherung sowie neuronale Speicherarchitekturen als physikalische Modelle eingeführt. Der Vergleich der Leistungsfähigkeit verschiedener Modellvarianten in unterschiedlichen Speicheraufgaben durch theoretische Analyse und Simulationsexperimente – der Gegenstand der folgenden Kapitel – setzt zunächst die Quantifizierung der verarbeiteten Information bei einer Speicheraufgabe voraus.

Wir geben hier zunächst eine kurze Einführung des Begriffs der Information in Daten und der Informationstheorie binärer Informationskanäle. Die Behandlung des Datenspeichers als Informationskanal führt am Ende dieses Kapitels zur Definition der Speicherkapazität als Gütemaß für Assoziativspeicher.

3.1 Zum Informationsbegriff

Aussagen oder Botschaften müssen, daß sie uns zu Ohren kommen oder ins Auge fallen können, in irgendeiner physikalischen Form vorliegen. Die physikalische Form einer Aussage nennt man ihre *Repräsentation*, die Übersetzung einer Aussage in ihre Repräsentation bezeichnet man als *Kodierung*. Repräsentationen von Aussagen können technisch gespeichert oder als Botschaften verschickt werden und in diesem Sinne spricht man von Datenverarbeitung. Wir beschränken uns hier auf *binäre Kodierungen*, welche Aussagen auf *binäre Muster* abbilden.

Betrachten wir zunächst Aussagen über den Zustand eines physikalischen Systems Σ. Die klassische Physik beschreibt den Systemzustand zu einem Zeitpunkt, indem die Eigenschaften des Systems auf den Wertebereich kontinuierlicher Variablen – der *Zustandsvariablen* von Σ – abgebildet werden. Trägt man den Wertebereich der Zustandsvariablen über den Zeitparameter auf, so erhält man den *Phasenraum* als Beschreibungsraum von Σ. Eine Aussage über den Zustand des Systems zum Zeitpunkt t (gewonnen durch eine Messung) entspricht einem Punkt im Phasenraum. Bei Kenntnis der physikalischen Gesetze des Systems[1] ist durch eine Messung das System vollständig determiniert, d. h. alle Aussagen über zukünftige und vergangene Systemzustände daraus ableitbar. Durch weitere Messungen kann keine zusätzliche Kenntnis über das System hinzugewonnen werden.

Eine ungenauere Beschreibung von Σ, etwa durch eine Rasterung des Phasenraums in eine endliche Anzahl disjunkter Zellen erlaubt i. a. keine vollständige Determinierung des Systems. Es können jedoch aufgrund einer Messung zum Zeitpunkt t immer noch probabilistische Vorhersagen getroffen werden, gegeben durch eine Wahrscheinlichkeitsverteilung auf

---

[1] Differentialgleichungen der Zustandsvariablen

den diskretisierten Zustandsvariablen von Σ [2].

Für alle nicht in der Zukunft liegenden Zeitpunkte sind durch weitere Messungen zusätzliche Aussagen über das System möglich, welche die Kenntnis über den Systemzustand weiter erhöhen[3].

Der Informationsgehalt solcher Aussagen bezüglich der bis dahin gültigen Vorhersage kann nun mathematisch quantifiziert werden. Sei Σ ein eindimensionales System, für welches wir probabilistische Vorhersagen kennen, d.h. die Wahrscheinlichkeitsverteilung einer diskreten Zufallsvariable x, deren Wertebereich $B_x := \{z_1, ..., z_m\}$ alle möglichen Zustände von Σ umfaßt. Für die Botschaft über einen eingetretenen Systemzustand „$x \in E$“, mit $E \subset B_x$, definiert man die *Information* der Botschaft als

$$I(E) := -\log_2[\, p[x \in E]] \,. \tag{3.1}$$

Sie stellt ein Maß dafür dar, wie schwer es ist, aufgrund der bekannten probabilistischen Beschreibung ihren Inhalt zu erraten. Wählt man den Logarithmus zur Basis Zwei, so ist $I(E)$ in der Einheit Bit gegeben und entspricht gerade der mittleren Anzahl von Ja-Nein Fragen, die zum geschickten Erfragen der Aussage notwendig sind.

Diese Definition ordnet auch jedem Systemzustand selbst, über die Mengen mit einem Element: $E = \{z\}$, einen Informationswert zu. Diese Zustandsgröße wird in der Physik als *Entropie* bezeichnet[4].

Für die Höhe der Information ist nicht die physikalische Form der Botschaft, die Repräsentation der Aussage E, sondern allein die ihr zugeordnete Vorhersagewahrscheinlichkeit relevant. Die Information einer Botschaft sinkt monoton mit der Wahrscheinlichkeit des Eintreffens der Aussage: umso wahrscheinlicher die faktisch gewordene Zustandbeschreibung z war, umso weniger werden sich die neuen, nach Erhalt der Botschaft gültigen Wahrscheinlichkeiten, von den bisherigen unterscheiden[5].

Die *Information in einer probabilistischen en Vorhersage* ist definiert als Summe der Erwartungswerte der Entropie aller darin möglichen Zustände

$$I(x) := -\sum_{z \in B_x} p[x = z] \log_2[\, p[x = z]] \leq \log_2[\sharp B_x] \,. \tag{3.2}$$

$\sharp B_x$ ist definiert als Anzahl der Elemente in B_x. Die Majorante wird für den Fall angenommen, daß die Zustandsvariable gleichverteilt ist. Die mittlere Information eines Vorhersagemodells steigt also mit wachsender Kardinalität des Wertebereichs von x, also je detaillierter die Wahlmöglichkeiten beschrieben werden und umso gleichmäßiger dieselben eintreffen.

3.2 Binärer Informationskanal

Binäre Zufallsvariable

Für den Fall, daß von einem System Σ eine probabilistische Vorhersage der binären Variable

---

[2] Die physikalische Vorhersage des Würfelwurfes für Zeitpunkte vor dem Stillstand des Würfels ist die Zustandsvariable eine gleichverteilte Zufallsvariable mit $p[x = i] = 1/6$. Die Gleichverteilung in der Vorhersage bedeutet immer soviel wie „man weiß, daß man nichts weiß“.

[3] Die zuvor gültigen Vorhersagen über mögliche Ereignisse werden durch in der neuen Messung festgestellte faktische Ereignisse schlagartig ihrer Gültigkeit beraubt. Der Gegenwartszeitpunkt ist dadurch, anders als beim deterministisch beschriebenen System besonders ausgezeichnet.

[4] Man beachte, daß die Definition der Entropie ganz und gar abhängig von der gewählten Diskretisierung des Phasenraums ist.

[5] Extrembeispiel: $I(B_x) = 0$, die Aussage $E = B_x$ ändert überhaupt nichts an den bestehenden Vorhersagewahrscheinlichkeiten.

$x \in \{a, 1\}$, durch die Verteilung $p := p[x = 1]$ beziehungsweise $p[x = 0] = 1 - p$ vorliegt, erhält man als mittlere Information der Zufallsvariablen x die *Shannon Information*

$$I(x) = -p\log_2[p] - (1-p)\log_2[1-p] =: i(p)\,. \tag{3.3}$$

Nun soll uns eine Botschaft, z. Bsp. des Inhalts „$x = 1$" überbracht werden, aber indirekt über einen Informationskanal. Der *binäre Informationskanal* bildet die Zufallsvariable (x,p) auf seine Ausgangsvariable $(\tilde{x},\tilde{p})$ ab, beide mit dem Wertebereich $B_x = B_{\tilde{x}} = \{a, 1\}$. Für den Kanalausgang müssen wir eine *Interpretation* kennen, eine Abbildung Γ, die jedem Ausgangszustand $\tilde{x} \in \tilde{Z}$ einen Zustand $x \in Z$ von Σ zuordnet[6].

Kanalrauschen

Im verrauschten Informationskanal wird ein Eingangszustand nicht in allen Fällen auf den via Γ korrespondierenden Zustand abgebildet, sondern kann durch kanalinterne Zufallsprozesse vertauscht werden. Die Eigenschaften des Kanals, werden beschrieben durch das Paar von Verwechslungswahrscheinlichkeiten $\tilde{\alpha}$ und $\tilde{\beta}$ in Abbildung 3.1. Hiermit ergibt sich die Verteilung

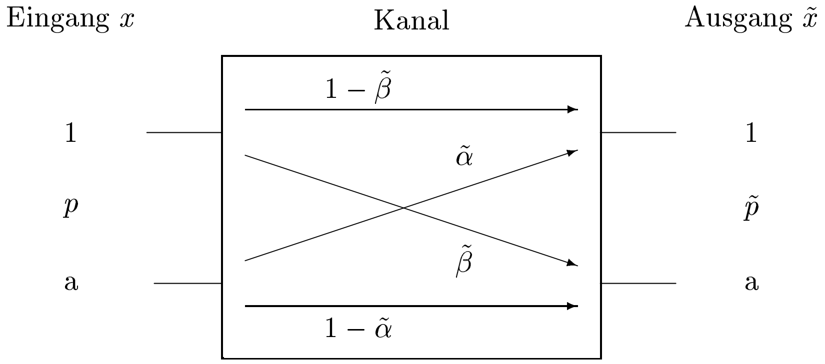

Abbildung 3.1: Binärer Übertragungskanal: Die Fehlerwahrscheinlichkeiten des Kanals sind als *bedingte Wahrscheinlichkeiten* definiert: $\tilde{\alpha} := p[\tilde{x} = 1 \mid x = a]$, $\tilde{\beta} := p[\tilde{x} = a \mid x = 1]$.

der Ausgangsvariable $\tilde{x}$ gemäß

$$\tilde{p} := p[\tilde{x} = 1] = p(1-\tilde{\beta}) + (1-p)\tilde{\alpha}. \tag{3.4}$$

Die *bedingte Information* einer durch den binären Kanal übertragenen Botschaft ist definiert als

$$\begin{aligned} I(x \mid \tilde{x}) &= \tilde{p}\, i(\, p[x = a \mid \tilde{x} = 1]) + (1-\tilde{p})\, i(\, p[x = 1 \mid \tilde{x} = a]) \\ &= [p(1-\tilde{\beta}) + (1-p)\tilde{\alpha}]\; i\left(\frac{(1-p)\tilde{\alpha}}{p(1-\tilde{\beta}) + (1-p)\tilde{\alpha}}\right) \\ &+ [p\tilde{\beta} + (1-p)(1-\tilde{\alpha})]\; i\left(\frac{p\tilde{\beta}}{p\tilde{\beta} + (1-p)(1-\tilde{\alpha})}\right) =: i\Big(p, \tilde{\alpha}, \tilde{\beta}\Big) \end{aligned} \tag{3.5}$$

Ausdruck (3.5) entspricht der mittleren Informationsmenge, die notwendig ist, um aus gegebenem $\tilde{x}$ den Wert von x zu rekonstruieren. Völlig unzuverlässe Übertragung mit $\tilde{p} = p[\tilde{x} =$

---

[6] Hier sind nur zwei verschiedene Abbildungen Γ möglich: Entweder ordnet Γ die identischen Werte zu oder vertauscht, man spricht dann vom invertierenden Kanal.

$1 \mid x = a]$, d.h. statistisch unkorrellierten Ein- und Ausgang, erhalten wir für $\tilde{\beta} = 1 - \tilde{\alpha}$. Es ist dann $i\left(p, \tilde{\alpha}, \tilde{\beta}\right) = i(p)$, wir brauchen zur Korrektur die volle Information in x. Für fehlerfreie Übertragung $\tilde{\beta} = \tilde{\alpha} = 0$, bzw. für konsequent invertierende Übertragung $\tilde{\beta} = \tilde{\alpha} = 1$ gilt jeweils $i\left(p, \tilde{\alpha}, \tilde{\beta}\right) = 0$. Schließlich ist die *Kanalkapazität* definiert als mittlere Transinformation

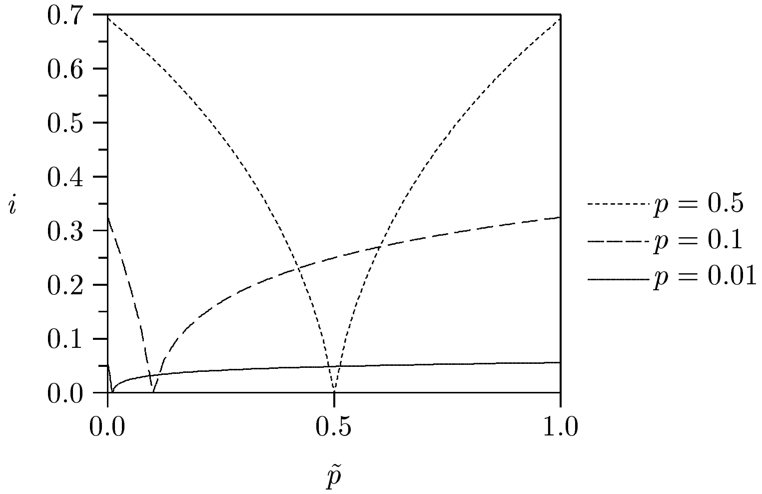

Abbildung 3.2: Bedingte Information $i(p, \tilde{\alpha}, \tilde{\beta})$ aufgetragen über die Ausgangsverteilung $\tilde{p}$ des Kanals, die sich bei fester Eingangsverteilung p durch Variation der Fehlerwahrscheinlichkeiten ändert. Jede der drei Kurven gehört zu einem festen Wert p. Es wurde immer für $\tilde{p} < p$ die Fehlerwahrscheinlichkeit $\tilde{\alpha} = 0$ gesetzt und für $\tilde{p} > p$ entsprechend $\tilde{\beta} = 0$.

über die Variable x die durch den Kanal übertragen wird

$$T(x, \tilde{x}) := i(p) - i\left(p, \tilde{\alpha}, \tilde{\beta}\right) =: t(p, \tilde{\alpha}, \tilde{\beta}) \tag{3.6}$$

3.3 Übertragung binärer Muster

Setzt man beim Transfer binärer Muster voraus, daß die Verrauschung in jeder Komponente unabhängig ist, kann der Informationskanal mit parallelgeschalteten binären Kanalkomponenten modelliert werden. Ein Einzelmuster wird übertragen, indem jeder binäre Aktivitätswert durch eine Kanalkomponente geschickt wird. Jede Kanalkomponente arbeitet wie in Abschnitt 3.2 beschrieben mit den gleichen Fehlerwahrscheinlichkeiten $\tilde{\alpha} := p[\tilde{x}_i = 1 \mid x_i = a]$ und $\tilde{\beta} := p[\tilde{x}_i = a \mid x_i = 1]$. Die bedingte Information eines Ausgangsmusters $\tilde{x}$ über x ist dann mit Gleichung (3.5) gegeben als

$$I(x \mid \tilde{x}) = n\, i\left(p, \tilde{\alpha}, \tilde{\beta}\right). \tag{3.7}$$

3.4 Zufallsgenerierung von Mustermengen

Ein binäres Muster der Dimension n ist festgelegt durch die Menge seiner On-Neurone. Ein Zufallsmuster x mit vorgegebener Aktivierung[7] wird nun alternativ mit einem der folgenden Zufallsauswahlprozesse generiert:

---

[7] siehe Abschnitt 2.2

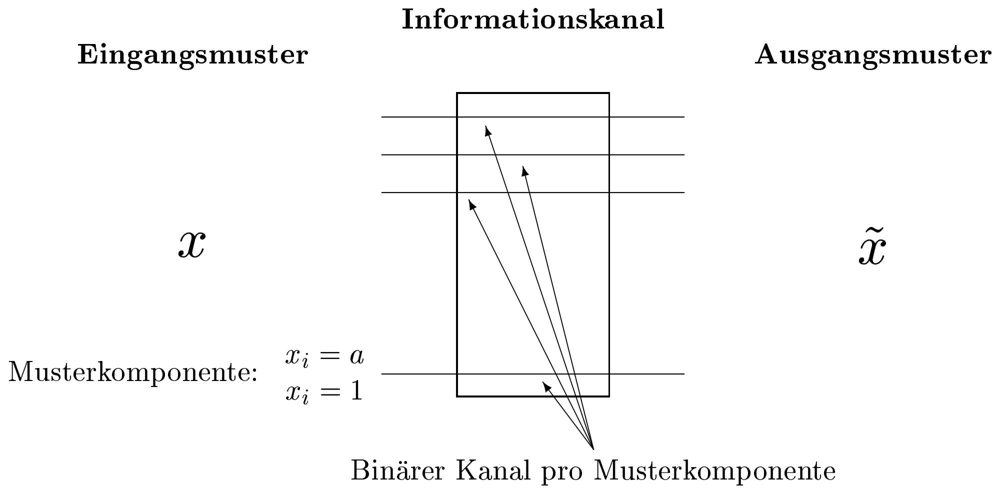

Abbildung 3.3: Der Informationskanal für ein n-dimensionales Muster besteht aus n binären Kanälen, jeder Kanal transportiert den Zustand eines Neurons. Kanalausgang i wird interpretiert als Aussage über den Zustand des i-ten Neurons im Eingangsmuster.

$Z(p)$ Für vorgegebene Aktivierung p wird jede Komponente eines Musters x in einem binären Zufallsprozess mit der Wahrscheinlichkeit p auf den Wert $x_i = 1$ gesetzt. Die Aktivierung eines auf diese Weise erzeugten Musters stimmt für endliche n nicht genau mit dem vorgegebenen Wert überein. Wegen dem Gesetz der großen Zahlen gilt jedoch für sehr große n für alle Muster: $p(x) = p$. Die mittlere Information eines so erzeugten Musters läßt sich mit (3.3) angeben als

$$I(x) = n\, i(p)\,. \tag{3.8}$$

$Z(k)$ Als schärfere Vorgabe wird anstelle der Aktivierung die genaue Anzahl der aktivierten Neurone k im Muster vorgeschrieben. Der Zufallsprozess bestimmt dann für jedes Muster eine Menge von On-Neuronen mit fester Kardinalität. Es gibt $\binom{n}{k}$ Wahlmöglichkeiten der Realisierung eines Musters aus denen mit Gleichverteilung gewählt wird. Die Information einer Realisierung, also eines Musters, beträgt gemäß (3.1)

$$I(x) \approx k\log_2 \left[\frac{n}{k}\right] \approx n\, i\left(\frac{k}{n}\right)\,. \tag{3.9}$$

Durch M-faches Ausführen von Z(p) oder Z(k) wird eine Menge voneinander unabhängiger Zufallsmuster erzeugt. Eine solche Menge von Zufallsmustern $\mathcal{S}$ ist durch drei Parameter charakterisiert: Dimension der Muster n, Anzahl der Muster M und die Aktivierung p beziehungsweise k/n für $Z(k)$. Für Heteroassoziation werden Adress- und Inhaltsmuster $\mathcal{S}^A$ und $\mathcal{S}^C$ voneinander unabhängig generiert, wir benutzen die Notation $\mathcal{S}^A =: \mathcal{S}(p, m, M)$ und $\mathcal{S}^C =: \mathcal{S}(q, n, M)$.

Die Eigenschaften eines mit $Z(p)$ generierten Zufallsmusters ist in Tabelle 3.1 dargestellt.

Bei Erzeugung mit $Z(k)$ enthält jedes Zufallsmuster genau k Einsen, für den Spezialfall $k = 1$ werden *singuläre Muster* mit genau einer On-Komponente erzeugt.

Für $k = np$ und falls $n >> k >> 1$ liefern beide Zufallsauswahlprozesse Muster mit dem selben Informationsgehalt. Der *mittlere Informationsgehalt* einer Menge $\mathcal{S}(q, n, M)$ von

	$n < \infty$	$n \to \infty$
Aktivierung	fluktuiert: $E(p(x)) = p$	konstant: $p(x) = p$
Anzahl v. Einsen	fluktuiert: $E(k) = np$	

Tabelle 3.1: Eigenschaften eines Einzelmusters aus einer mit Z(p) erzeugten Menge von n-dimensionalen Zufallsmuster.

Zufallsmustern ist mit (3.8) bzw. mit (3.9) ungefähr

$$I(\mathcal{S}) = Mn\,i(q)\,.$$

3.5 Übertragung von Mustermengen

Um Speicherung und assoziative Abfrage eines Einzelmusters im Speicher zu beschreiben, modellieren wir den Transfer einer Menge von Zufallsmustern $\mathcal{S}(n, M, p)$ als serielle Übertragung von Einzelmustern. Die bedingte Information erhält man dann als Summe der bedingten Information aller M Musterpaare

$$I(\mathcal{S} \mid \tilde{\mathcal{S}}) = \sum_{i=1}^{M} I\Big(x^i \mid \tilde{x}^i\Big) = Mn\,i\Big(q, \tilde{\alpha}, \tilde{\beta}\Big)\,. \tag{3.10}$$

Die Kanalkapazität ist dann wieder wie in (3.6) gegeben durch

$$T(\mathcal{S}, \tilde{\mathcal{S}}) = I(\mathcal{S}) - I(\mathcal{S} \mid \tilde{\mathcal{S}}). \tag{3.11}$$

3.6 Speicherkapazität

Als Inhaltsmuster $\mathcal{S}^C$ und Adressmuster $\mathcal{S}^A$ werden Mengen von Zufallsmuster gleicher Aktivierung gemäß Abschnitt 3.4 generiert und für das Lernverfahren benutzt.

3.6.1 Retrievaltestschema

Für den Retrievalprozess führen wir folgendes Testschema ein:

- **Adressierung:** Durch einen verrauschten Kanal B (Benutzerkanal) wird das gegebenenfalls unscharfe Wissen eines Benutzers über die Adressen modelliert. B erzeugt durch Vorgabe der Fehlerwahrscheinlichkeiten $\hat{\alpha}$ und $\hat{\beta}$ aus den Adressmustern eine Menge von verrauschten *Testadressierungsmustern* $\hat{\mathcal{S}}^A$. Der Fall *perfekter Adressierung*: $\hat{\alpha} = \hat{\beta} = 0$ entspricht vollständiger Kenntnis des Benutzers über die Adressen.
- **Retrieval:** Durch Abfrage mit den Testadressierungsmuster wird aus dem Speicher die Menge der Ausgangsmuster $\tilde{\mathcal{S}}^C$ ausgelesen. Die Reproduktion der Inhaltsmuster $\mathcal{S}^C$ am Speicherausgang wird durch den verrauschten Kanal D (für Datenkanal) beschreiben. Die Fehlerwahrscheinlichkeiten $\tilde{\alpha}$ und $\tilde{\beta}$, die von der Adressierung abhängen,

werden entweder durch Experimente oder durch Analyse der Speichermodelle bestimmt. Dieser Kanal setzt sich aus zwei Abschnitten zusammen (vgl. Abbildung 3.4): Erster Abschnitt bildet der Speicherprozess, der die Lernmuster $\mathcal{S}^A$ und $\mathcal{S}^C$ in die Synapsenmatrix transformiert. Der Retrievalvorgang als zweiter Abschnitt transformiert die synaptische Kopplungsmatrix zur Menge von Adressmuster $\tilde{\mathcal{S}}^C$.

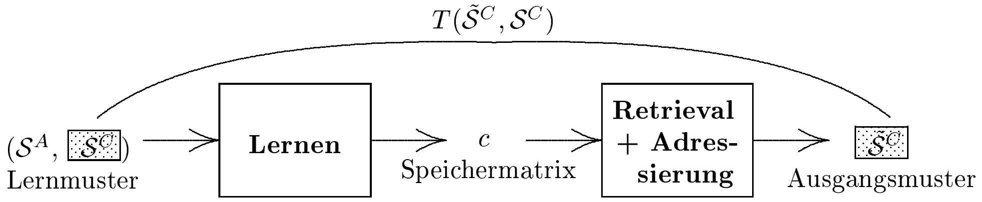

Abbildung 3.4: Dateninformationskanal D: Linker Abschnitt: Lernvorgang; Rechter Abschnitt: Retrievalvorgang.

3.6.2 Kapazitätsdefinitionen

Wohl jede Anwendung eines Speichers gibt obere Grenzen an die Retrievalfehlerwahrscheinlichkeiten vor, die nicht überschritten werden dürfen. Wir nennen einen Satz solcher Grenzen für beide Fehlertypen ein *Qualitätskriterium.* Ein Gütemaß für Retrieval sollte für vorgegebenes Qualitätskriterium die maximal speicherbare Information messen und sie in Relation zum Platzbedarf des Speichermediums, im Falle neuronaler Modelle einfach zur Anzahl der Synapsen setzen. Um das Maximum auszuloten, wird die Anzahl der gespeicherten Muster M auf den maximalen Wert M^* gesetzt, der die Einhaltung des Qualitätskriterium noch gewährleistet. M^* ist eine von der Adressierung abhängige Größe, sie muß mit zunehmenden Adressierungsfehlern wegen der damit ansteigenden Retrievalfehler reduziert werden. Es folgen drei verschiedene Definitionen für Speicherkapazität:

- Die *Musterkapazität* mißt den Informationsgehalt der Inhaltsmuster pro Synapse:

$$P(m,n) := \max_M \{I(\mathcal{S}^C)\}/nm = M^*\, i(q)\,/m. \tag{3.12}$$

- Die *Ausgangskapazität* ist definiert als die Kanalkapazität des Datenkanals D in Abbildung 3.4 pro Synapse:

$$\begin{aligned} A(m,n) &:= \max_M \left\{T(\mathcal{S}^C, \tilde{\mathcal{S}}^C)\right\}/mn = P(m,n) - \frac{M^*}{m} I(y_i^k \mid \tilde{y}_i^k) \\ &= \frac{M^*}{m}\left[i(q) - i\Big(q, \tilde{\alpha}, \tilde{\beta}\Big)\right] \end{aligned} \tag{3.13}$$

- Für Autoassoziation betrachten wir die Differenz der Transinformation über die Inhaltsmuster, die einerseits in den Adressmustern, andererseits in den Ausgangsmustern enthalten ist. Diese Informationsdifferenz, wieder normiert auf eine Synapse, definiert die *Korrekturkapazität*:

$$\begin{aligned} K(n) &:= \max_{\hat{\mathcal{S}}^C} \left\{T(\tilde{\mathcal{S}}, \mathcal{S}) - T(\hat{\mathcal{S}}, \mathcal{S})\right\}/n^2 \\ &= \max_{\hat{\alpha},\hat{\beta}} \frac{M^*}{n}\left[i\Big(q, \hat{\alpha}, \hat{\beta}\Big) - i\Big(q, \tilde{\alpha}, \tilde{\beta}\Big)\right] \end{aligned} \tag{3.14}$$

Während die Musterkapazität den Informationsverlust durch Retrievalfehler in $\tilde{\mathcal{S}}$ unberücksichtigt läßt, berücksichtigt die Ausgangskapazität die Eigenschaften der wirklich vom Speicher gelieferten Daten. Beide Größen stimmen überein, falls das Qualitätskriterium für hinreichend kleine Fehler im Ausgangsmuster sorgt, so daß der zweite Term in Gleichung 3.13 verschwindet. Das Maximum erreichen diese beiden Größen immer für perfekte Adressierung, um Fehlertoleranz des Retrievals zu beurteilen, muß für verrauschte Adressierung gemessen werden.

Die Korrekturkapazität berücksichtigt bei Autoassoziation die Bilanz der Kanalkapazitäten von Daten- und Benutzerkanal, siehe Abbildung 3.5. Sie stellt ein geeignetes Gütemaß für Mustervervollständigung dar[8] und wird in Kapitel 6 als Gütekriterium im Vergleich der unterschiedlichen Modellvarianten benutzt.

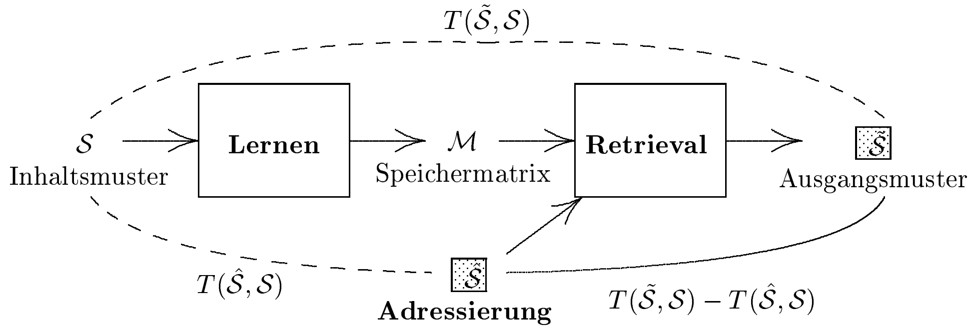

Abbildung 3.5: Korrekturkapazität: Informationsbilanz von Datenkanal D und Benutzerkanal B bei Autoassoziation.

In [Pa92] wird außerdem noch die *Lernkapazität* eingeführt. Sie ist definiert als Kanalkapazität des Lernvorgangs, also von dem linken Abschnitt des Datenkanals in Abbildung 3.4. Als Kapazität eines Kanalabschnitts stellt die Lernkapazität eine obere Schranke der Kapazität des Gesamtkanals dar. Es wird in [Pa92] gezeigt, daß für die Lernkapazitäten bei autoassoziativem Lernen L_A und heteroassoziativem Lernen L_H immer die Beziehung gilt: $L_H = 2L_A$. Diese Relation ist iintuitiv plausibel, da autoassoziatives Lernen für lokale Lernregeln immer zu einer symmetrischen Konnektivitätsmatrix führt, bei der jeweils zwei Matrixelemente identisch sind. Bei Autoassoziation ist daher das zur Verfügung gestellte Speichermedium nur zur Hälfte ausnutzbar (oder die Synapsenmatrix asymptotisch sehr leicht auf die halbe Zahl von Werten komprimierbar).

3.6.3 Hifi-Qualitätskriterium

Eine wichtige Rolle werden in dieser Arbeit *spärliche Muster* spielen, die ein sehr kleines Verhältnis zwischen On- und Off-Komponenten besitzen: $p = k/n << 1$. Für $n \to \infty$ sind nichttriviale Muster mit *asymptotischer Spärlichkeit* möglich, mit $p \to 0$ aber $np \to \infty$. Interessanterweise ist es gerade die Speicherung solcher Muster, die sich als optimal bezüglich der Speicherkapazität herausstellt.

Um die Ähnlichkeit zweier spärlicher Muster zu gewährleisten, reicht es nicht, kleine absolute Fehlerwahrscheinlichkeiten zu fordern, vielmehr müssen die Fehleranteile in On- und

---

[8] so erreicht sie ihr Maximum für verrauschte Adressierung, also in einem Bereich, in dem wirklich Mustervervollständigung stattfindet.

Off-Neuronen des Inhaltsmuster klein werden, d.h. die *Fehlerverhältnisse*:

$$\tilde{g} := \tilde{\beta}/(1-p) \tag{3.15}$$

$$\tilde{f} := \tilde{\alpha}/p \tag{3.16}$$

Mit der Schwellenwahl am Assoziativspeicher wird die Verteilung der Retrievalfehler auf On- und Off-Komponenten gesteuert. Für Muster mit $p = 0.5$ sollte $\tilde{\alpha} = \tilde{\beta}$ angestrebt werden, dagegen sollte für spärliche Muster $\tilde{f} = \tilde{g}$ eingehalten werden.

Das *Hifi-Qualitätskriterium* fordert nun gerade das asymptotische Verschwinden des Korrekturterms in der Ausgangskapazität (und damit ihre Übereinstimmung zur Musterkapazität)

$$I\left(\mathcal{S} \mid \tilde{\mathcal{S}}\right)/nm \to 0 \text{ für } n, m \to \infty. \tag{3.17}$$

Eine notwendige Bedingung hierfür ist das Verschwinden der Fehlerverhältnise im Ausgangsmuster

$$\tilde{g} \to 0 \text{ und } \tilde{f} \to 0 \text{ für } n, m \to \infty. \tag{3.18}$$

Die *Hifi-Kapazitäten* erhält man aus den Definitionen in Abschnitt 3.6.2 für $n, m \to \infty$, wenn die Anzahl der Inhaltsmuster M^* unter Einhaltung des Hifi-Qualitätskriterium optimiert wird, d.h.:

$$M^* = \max\left\{M : \frac{M}{m} I(y_i^k \mid \tilde{y}_i^k) \to 0\right\}. \tag{3.19}$$

Kapitel 4

Einschrittretrieval

4.1 Statistische Analyse

Gemäß dem Testschema aus Kap. 3.9.1 gehen wir in diesem Kapitel davon aus, daß der Lernprozess aus zufallsgenerierten Adress- und Inhaltsmuster $(\mathcal{S}^A, \mathcal{S}^C)$ die Speichermatrix gebildet hat und abgeschlossen ist.

Für Einschrittretrieval ist der Assoziativspeicher ein reines feed-forward System ohne Rückkopplung, bei dem ein anliegendes Eingangsmuster ein bestimmtes Ausgangsmuster erzeugt. Jedes Neuron j „entscheidet" im Updateprozess, abhängig vom jeweiligen dendritischen Potential d_j (2.1) und dem Schwellenwert $\Theta_j =: \Theta$, ob es feuert oder passiv bleibt. Dieser Vorgang kann als Signal-Detektionsproblem des Neurons betrachtet werden, wobei die Zufallsvariable d_j die Rolle des Detektionssignals spielt und Θ die Rolle der Detektionsschwelle.

In unserem Modell kann der Benutzer einen Wert Θ wählen, auf den die Schwelle uniform für jedes Neuron gesetzt wird. Der Wert Θ kann zum Beispiel abhängig von der Aktivierung des Eingangsmuster gewählt werden, oder er kann zur Regelung der Aktivierung des Ausgangsmusters auf eine vorgegebene Größe benutzt werden. Für uniforme Schwellenwahl ist es möglich, die Signal-Detektionsprobleme für die beiden Kollektive der On- und Off-Neurone zusammenzufassen. Das kollektive Neuronenverhalten wird auf ein repräsentatives Neuron in einem mittleren dendritisches Potential reduziert[1]. Für die gemittelten dendritischen Potentiale verwenden wir die Notation:

$$d_{on} \quad := \quad < d_j >_{j \in \{j : y_j^k = 1\}} \tag{4.1}$$

$$d_{off} \quad := \quad < d_j >_{j \in \{j : y_j^k = a\}} \tag{4.2}$$

Mit Annahmen über die Eingangsmuster, wie sie Abschnitt 4.3 festlegt, können diese Größen als Zufallsvariablen behandelt werden. Es ist dann mit elementaren statistischen Methoden möglich, die Retrievalfehlerwahrscheinlichkeiten

$$\tilde{\beta} \quad = \quad p[\tilde{x}_i = a \mid x_i = 1] = \; p[d_{on} < \Theta] \tag{4.3}$$

$$\tilde{\alpha} \quad = \quad p[\tilde{x}_i = 1 \mid x_i = a] = \; p[d_{off} \geq \Theta] \tag{4.4}$$

für Systeme endlicher Größe und für Systeme divergierender Systemgröße[2] zu bestimmen. Mit den Fehlerwahrscheinlichkeiten $\hat{\alpha}$ und $\hat{\beta}$ lassen sich die in Abschnitt 3.6.2 definierten Speicherkapazitäten berechnen.

---

[1] Diese Vorgehensweise wird in der Physik als Mean-Field Nährung bezeichnet, siehe Abschnitt 7.1.2.

[2] Dieser Fall wird in der Physik als thermodynamischer Limes bezeichnet, wir sprechen hier von der *asymptotischen Rechnung.*

4.2 Signal-zu-Rausch-Verhältnis

Als schwellenunabhängiges Qualitätskriterium für das Retrieval führen wir das *Signal-zu-Rausch-Verhältnis* ein. Bezüglich eines Inhaltsmusters zerlegen wir das dendritische Potential (4.1) in Signal- und Rauschkomponente

$$d_{on} =: s_{on} + N_{on}\ , \quad d_{off} =: s_{off} + N_{off} \tag{4.5}$$

und definieren das Signal-zu-Rausch-Verhältnis als

$$r := E(s_{on} - s_{off})/\sigma(N). \tag{4.6}$$

Mit steigendem r verbessert sich die Retrievalqualität. Für eine bestimmte Schwellenwahl Θ werden umso mehr Neurone ihr Signal-Detektionproblem richtig lösen, umso größer $E(s_{on} - s_{off})$ und umso kleiner $\sigma(N)$. Die Verwendung des Signal-zu-Rausch-Verhältnis (4.6) als Qualitätskriterium ist gleichbedeutend mit den Annahmen, daß erstens $E(N_{on}) = E(N_{off})$ gilt und zweitens, daß die Varianz von N die entscheidende Größe für die Retrievalfehler ist. Aus dem Signal-zu-Rausch-Verhältnis erhält man für einen Schwellenwert auch wieder leicht die Fehlerwahrscheinlichkeiten ,vgl. (4.31) und (4.32).

4.3 Adressierung

4.3.1 Festgehaltene Musteraktivierung

Für vorgegebene Aktivierung p bzw. q werden Adress- und Inhaltsmuster mit Methode $Z(p)$ aus Abschnitt 3.4 zufallserzeugt. Die Adressierung ist durch die Fehlerwahrscheinlichkeiten

$$\hat{\alpha} = \ p[\hat{x} = 1 \mid x = a]\ , \quad \hat{\beta} = \ p[\hat{x} = a \mid x = 1] \tag{4.7}$$

des B Kanals festgelegt, der aus den Adressmustern die Eingangsmuster $\hat{\mathcal{S}}^A$ für das Retrieval erzeugt.

Wir behandeln hier das Signal-Detektionsproblem für drei unterschiedliche Fälle:

a) Fehlerfreies Adressmuster x^ν mit der Aktivität $\hat{k} := \sharp\{i : x_i^k = 1\}$ als Eingangsmuster.

b) Ein *Ensemble von Retrievalprozessen* mit fehlerfreien Adressmustern mit der Aktivierung $p = k/m$ als Eingangsmuster.

c) Ein *Ensemble von Retrievalprozessen mit verrauschter Adressierung.* Es werden hier Eingangsmuster mit $\hat{\alpha} = 0$ benutzt, die Eingangsmuster werden also ausschließlich durch On-Fehler, also fehlende Einsen, verrauscht. Die Aktivierung eines Eingangsmusters hat damit den Wert $\hat{p} = p(1 - \hat{\beta})$.

Für die Fälle b) und c) erhalten wir die über das Ensemble der Retrievalprozesse gemittelten Fehler. Die Eingangsaktivierung für jeden Retrievalprozess im Ensemble wird für großes m wegen dem Gesetz der großen Zahlen mit der Aktivierung $\hat{p}$ übereinstimmen. In der weiteren Analyse spielt die *Eingangsaktivität* des Retrievalprozesses eine wichtige Rolle. Sie ist nach (2.4) gegeben als

$$\hat{\mu} = \mu(\hat{\beta}) := p(1 - \hat{\beta}) + \Big(1 - p(1 - \hat{\beta})\Big)\, a \tag{4.8}$$

und vereinfacht sich für perfekte Adressierung zu

$$\mu := \mu(0) = p + (1 - p)a. \tag{4.9}$$

4.3.2 Feste Anzahl der On-Neurone

Wir werden noch eine zweite Variante der Lernmustergenerierung und Adressierung untersuchen. Dabei werden in jedem Retrievalschritt des Ensembles die Anzahlen der On-Neurone von Adress-, Inhalts- und Eingangsmuster konstant gehalten, sie waren in Abschnitt 4.3.1 binomiale Zufallsvariable. Die Zufallserzeugung der Adress- und Inhaltsmuster für feste Anzahlen von On-Neuronen k bzw. l der Muster erfolgt mit Methode $Z(k)$, das Eingangsmuster wird dadurch verrauscht, indem im Adressmuster g_1 On-Neurone den Wert a gesetzt werden, also g_1 On-Fehler erzeugt werden und entsprechend g_2 Off-Fehler, wobei die Positionen der Fehler zufällig gewählt werden. Das Eingangsmuster enthält also stets

$$\hat{k} = k + g_1 - g_2 \tag{4.10}$$

Einsen, für die Aktivität gelten auch die Gleichungen (4.9) und (4.8) mit $p = k/m$ und $\hat{\beta} = g_2/(m-k)$.

4.4 Binäre Speicherung mit der Hebbregel

Bei binärer Speicherung liegen die Einträge der Matrix in der Menge $\{0,1\}$. Wir benutzen $a = 0$, wodurch man ein maximales Signal-zu-Rausch-Verhältnis erhält, weil $E(s_{on} - s_{off})$ von a invariant ist, aber $\sigma(N)$ minimal wird für $a = 0$.

4.4.1 Fehlerwahrscheinlichkeiten

Adressiert man ein Inhaltsmuster y^ν, dessen Adressmuster x^ν k On-Neurone besitzt, durch ein Eingangsmuster mit g_1 On-Fehlern und g_2 Off-Fehlern, dann erhält man im Retrieval die Fehlerwahrscheinlichkeiten (4.4) und (4.3)

$$\tilde{\alpha} = \begin{cases} 1 & \text{für } \Theta = 0 \\ \sum_{l=\Theta}^{\hat{k}} \binom{\hat{k}}{l}\, p[\cap_{i=1}^{l}(\bar{c}_{ij} = 1) \cap_{i=l+1}^{\hat{k}} (\bar{c}_{ij} = 0)] & \text{für } \Theta \leq \hat{k} \\ 0 & \text{für } \Theta > \hat{k} \end{cases} \tag{4.11}$$

$$\tilde{\beta} = \begin{cases} 1 & \text{für } \Theta > \hat{k} \\ \sum_{l=1}^{\hat{k}-\Theta} \binom{g_1}{l}\, p[\cap_{i=1}^{l}(\bar{c}_{ij} = 1) \cap_{i=l+1}^{\hat{k}-\Theta} (\bar{c}_{ij} = 0)] & \text{für } \Theta > \hat{k} - g_2 \\ 0 & \text{für } \Theta \leq \hat{k} - g_2 \end{cases} \tag{4.12}$$

Der Ausdruck $p[\cap_{i=1}^{l}(\bar{c}_{ij} = 1) \cap_{i=l+1}^{\hat{k}} (\bar{c}_{ij} = 0)]$ steht für die Wahrscheinlichkeit mit der aus $\hat{k}$ Matrixelementen einer Off-Spalte des Retrievalmusters genau l Stück durch die Speicherung der $M' := M - 1$ anderen Muster auf Eins gesetzt werden. Die durch den Lernprozess von Zufallsmustern verursachten statistischen Abhängigkeiten verschiedender Matrixelemente ist Gegenstand der Propositionen in Anhang A. Proposition 1 liefert für den Fall festgehaltener Aktivierung die Wahrscheinlichkeit, daß ein Matrixelement unverändert bleibt:

$$p_0 := p[\bar{c}_{ij} = 0] = (1 - pq)^{M'}. \tag{4.13}$$

Zwei Elemente einer Spalte der synaptischen Kopplungsmatrix sind durch den Speichervorgang korreliert, ihre Kovarianz ist

$$\mathrm{Cov}[\bar{c}_{1j}\bar{c}_{2j}] = p_0^2 M p^2 q. \tag{4.14}$$

Proposition 2 behandelt den Fall der festgehaltenen Anzahl von Einsen. Für große Systemgrößen stimmen die Ergebnisse mit denen von Proposition 1 überein, für kleine Systeme ist p_0 größer und $\mathrm{Cov}[\bar{c}_{1j}\bar{c}_{2j}]$ kleiner.

Die folgende Abschätzung nach Proposition 4 in (4.11) und (4.12) (aus Proposition 1) berücksichtigt die synaptischen Korrelationen innerhalb einer Matrixspalte

$$p[\cap_{i=1}^{l}(\bar{c}_{ij}=1)\cap_{i=l+1}^{k}(\bar{c}_{ij}=0)] \tag{4.15}$$

$$\approx \left\{1+\sum_{h=1}^{l}\binom{l}{h}(-1)^h \exp\left[-M'p(1-p)^j\left\{1-(1-p)^h\right\}\right]\right\} \exp\left[-(M-l)p\left\{1-(1-p)^j\right\}\right]$$

mit $M' = M - 1$ u.s.w. Diese Berechnung der Retrievalfehler unter Berücksichtigung der Abhängigkeiten verschiedener Synapsen wird in Kapitel 6 benutzt.

4.4.2 Vernachlässigung synaptischer Korrelationen

Der Kovarianzterm (4.14) wird für spärliche Muster klein. Proposition 3 aus Anhang A zeigt für den asymptotischen Fall $m \to \infty$, daß für spärliche Adressmuster mit der Eigenschaft $m^{2/3}p \to 0$ die Matrixeinträge als unabhängig angenommen werden können. Auch im endlichen System erhalten wir für spärliche Muster bereits gute Fehlervorhersagen, wenn wir die Spalteneinträge als unabhängig annehmen.

Mit dieser Annahme erhält man in (4.11) und (4.12) Binomialverteilungen. Da Binomialsummen schwer handhabbar sind, benutzt man dafür gerne Nährungen, wie die Gauss- oder Poissonverteilung. Leider ist jedoch gerade der Bereich extrem spärlicher Muster (z.Bsp. mit $k = 10$ und $n = 4096$) der sich als besonders interessant herausstellen wird, mit keiner dieser Näherungen zugänglich.
Die drei Adressierungsfälle aus Abschnitt 4.3 führen zu folgenden Fehlerwahrscheinlichkeiten:

a) Perfekte Adressierung ergibt an einem On-Neuron das dendritische Potential von $d_{on} = \hat{k}$. Somit ist die maximale Schwellenwahl, die noch $\tilde{\beta} = 0$ garantiert: $\Theta = \hat{k}$. Man erhält damit aus (4.11).

$$\tilde{\alpha} = (1 - p_0)^k \tag{4.16}$$

b) Mittelt man über ein Ensemble von Retrievalprozessen, bei denen der Schwellenwert jeweils individuell auf $\Theta = k$ für jedes Eingangsmuster gesetzt wird, so ist die Schwelle eine fluktuierende Größe. Wir betrachten anstelle dessen eine mittlere Schwellenwahl, für jedes Eingangsmuster wird der Wert $\Theta = E\hat{k}$ benutzt. Für diese Schwellenwahl kann dann einfach der Mittelwert in (4.16) eingesetzt werden

$$\tilde{\alpha} = (1 - p_0)^{mp}. \tag{4.17}$$

Für den individuellen Schwellenabgleich ist zu beachten, daß wegen den Eingangsmustern mit $k < E\hat{k}$ gilt: $\tilde{\beta} > 0$. Außerdem wird $E\tilde{\alpha}(\Theta) > \tilde{\alpha}(E\Theta)$ sein, weil $\tilde{\alpha}(\Theta)$ eine konkave Funktion ist.

c) Für das Ensemble von Retrievalprozessen mit verrauschter Adressierung erhalten wir bei der Schwellenwahl $\Theta = (1 - \hat{\beta})E(\hat{k})$ die Fehlerwahrscheinlichkeit

$$\tilde{\alpha}(\hat{\beta}) = (\tilde{\alpha})^{(1-\hat{\beta})}, \tag{4.18}$$

vgl. Abbildung 6.1.

4.4.3 Vergleich unterschiedlicher Adressfehler

Betrachten wir nun das Zahlenbeispiel $k = 10$ und $p_0 = 0.5$. Wenn im Eingangsmuster eine Eins fehlt, erhalten wir mit $\Theta = 9$ den Fehler $\tilde{\alpha} = (0.5)^9$ gemäß (4.17). Für eine Adressierung mit $\hat{\beta} = 0$, aber einer zusätzlichen Eins, erhalten wir mit der Schwellenwahl $\Theta = k$ in (4.11)

$$\tilde{\alpha} = \sum_{i=k}^{\hat{k}} \binom{\hat{k}}{i} (1-p_0)^i p_0^{\hat{k}-i} = (k+1)(0.5)^{11} = 2.75(0.5)^9,$$

der Fehler ist also fast dreimal so hoch.

Für nichtallzukleines k (hier $k > 3$) wirkt sich in der Adressierung Noise durch zusätzliche Einsen ungünstiger auf die Retrievalfehler aus im Vergleich zu Noise durch fehlende Einsen.

4.4.4 Heteroassoziation

Die Wahrscheinlichkeit p_0 aus (4.13) sinkt, wenn die Anzahl der Inhaltsmuster erhöht wird. Da offensichtlich aus einer synaptischen Matrix mit uniformen Elementen keine Information gezogen werden kann, schließen wir im weiteren die Extremfälle des vollständig leeren Speichers $p_0 = 1$ und des vollständig überladenen Speichers $p_0 = 0$ aus.

Die obere Grenze der Anzahl speicherbarer Muster M^* für vorgegebene Fehlerschranke $\tilde{\alpha}$ ergibt sich aus den Beziehungen (4.13) und (4.16) als

$$M^* = \frac{\ln[p_0]}{\ln[1-pq]} = \frac{\ln\left[1-(\tilde{\alpha})^{1/mp}\right]}{\ln[1-pq]}. \tag{4.19}$$

Für $\tilde{\beta} = 0$ und $\tilde{f} = \tilde{\alpha}/q << 1$ vereinfacht sich (3.13), wir erhalten die Ausgangskapazität zu

$$A(m,n) \approx (M^*/m)\left\{i(q) - q(1-q)\tilde{f}\log_2\left[\tilde{f}(1-q)\right]\right\}. \tag{4.20}$$

Abbildung 4.1 zeigt links das Verhältnis von Musterzahl zu Neuronenanzahl M^*/m und rechts die Ausgangskapazität für verschiedene endliche Speichergrößen, $q = p$ und konstantes (kleines) Fehlerverhältnis. Für jede Systemgröße werden die Maxima für p-Werte angenommen, die proportional zu $\log[n]/n$ sind. Einen kleinen Vorgriff auf Simulationsexperimente stellt Abbildung 4.2 dar. Dort sind die Fehlerprognosen der vereinfachten Theorie mit Daten eines Experiments verglichen und zwar genau für die Anzahl von Inhaltsmustern aus Abbildung 4.1. Bei idealer Übereinstimmung würde man eine Waagerechte erwarten. Im Bereich der Maxima stimmen bereits bei den kleinen Systemgrößen die von der vereinfachten Theorie vorhergesagten Retrievalfehler gut überein.

Abbildung 4.2 gibt also für endliche Systeme in dem Parameterbereich ein realistisches Bild, in dem der assoziative Speicher die höchste Kapazität erreicht: für spärliche Inhaltsmuster mit $k \propto \ln[n]$.

Hifi-Ausgangskapazität

Nichtverschwindende asymptotische Ausgangskapazität ist nur bei $M^*/m > 0$ für $m \to \infty$ zu erzielen. In Gleichung (4.19) kann dies, da wir $p_0 \to 0$ ausgeschlossen haben, nur für $pq \to 0$ erreicht werden. Es ist dann

$$M^* \approx \frac{\ln[p_0]}{-pq}. \tag{4.21}$$

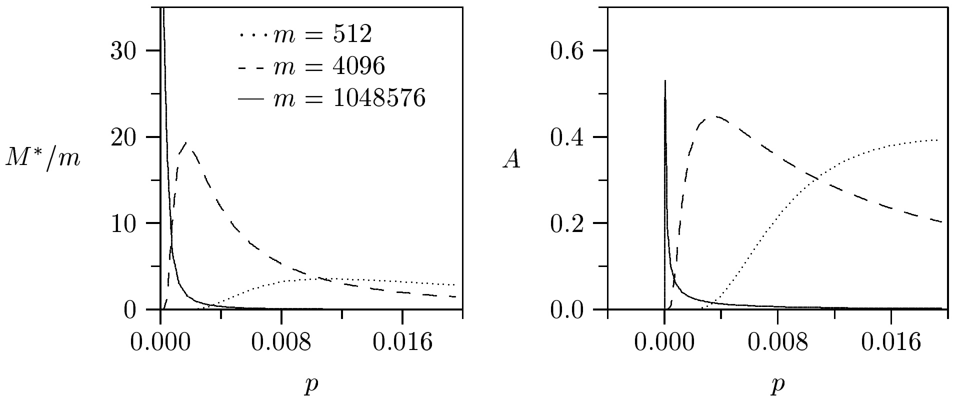

Abbildung 4.1: Maximale Anzahl speicherbarer Muster M^* und Ausgangsskapazität A in Bit/Synapse für ein vorgegebenes Fehlerverhältnis $\tilde{f} = \tilde{\alpha}/p = 0.01$. Binärer Speicherprozess, verschiedene endliche Speichergrößen, für den Fall $p = q$ and $n = m$.

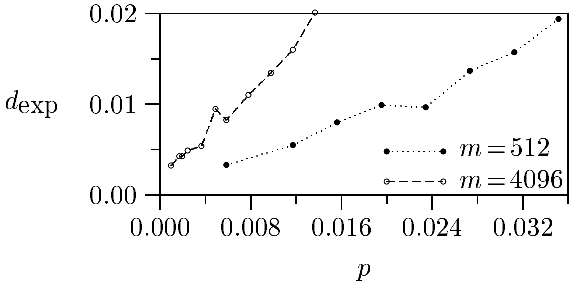

Abbildung 4.2: Retrievalfehlerverhältnis $\tilde{f}_{exp} = \tilde{\alpha}/p$ von Simulationen entlang der A-p Kurven von Abbildung 4.1, d.h. für $\tilde{f}_{theo} = 0.01$. Die Abweichung der Theorie hat zwei Gründe: i) In der Theorie wurde die Aktivierung der Lernmuster festgehalten, im Experiment die Anzahl der Mustereinsen festgehalten. Hierdurch liegt für kleine p Werte der experimentelle Fehler unter dem Theoriewert (Vergleich in Anhang A von Prop. 1 mit Prop. 2). ii) Für größere p Werte ist der theoretische Wert zu klein, weil dann die statistische Abhängigkeit zwischen verschiedenen Matrixelementen nicht mehr vernachlässigt werden darf.

Das Hifi-Kriterium führt mit der Beziehung (4.17) auf die folgende Bedingung an p und q:

$$\tilde{\alpha}/q = \exp[mp \ln[1-p_0] - \ln[q]] \to 0. \tag{4.22}$$

Falls $q \to 0$, dann kann (4.22) durch folgenden Ansatz erfüllt werden

$$p = u \frac{-\ln[q]}{m}, \tag{4.23}$$

wobei die positive Konstante u eine untere Schranke nicht unterschreiten darf

$$u > -(\ln[1-p_0])^{-1}. \tag{4.24}$$

Setzen wir (4.23) in (4.21) ein, dann erhalten wir die Ungleichung

$$M^* < \frac{m \ln[p_0] \ln[1-p_0]}{-q \ln[q]}, \tag{4.25}$$

welche in Gleichung (4.20) für $p_0 = 0.5$ und $m \to \infty$ die maximale Hifi-Ausgangskapazität von $A \approx 0.69$ Bit/Synapse liefert. Das Optimum wird immer bei größtmöglicher Spärlichkeit der Inhaltsmuster erreicht, unter der gerade noch das Hifi-Qualitätskriterium erfüllbar ist.

- **Bemerkungen:**

i) Für singuläre Adressmuster und beliebiges $q =$ const ist fehlerfreies Retrieval möglich, solange die kombinatorische Beschränkung für nichtüberlappende singuläre Muster $M^* \leq m$ eingehalten wird. Man erhält in diesem Fall mit (4.20) eine Ausgangskapazität von $A = i(q) \leq 1$ Bit/Synapse. Trotz des maximal möglichen Wertes der Ausgangskapazität ist dieser Fall uninteressant. Da weder verteilte Speicherung noch Fehlertoleranz gegeben ist, besteht kein Vorteil der neuronalen Implementation gegenüber herkömmlichen Speicherbausteinen.

ii) Die Ausgangskapazität verschwindet einerseits für konstantes p, denn dann fordert der Ansatz (4.23) asymptotisch leere Inhaltsmuster $q \propto \exp[-mp/u]$ und andererseits für singuläre Inhaltsmuster, wieder wegen der kombinatorischen Beschränkung $M^* \leq m$.

4.4.5 Autoassoziation

Für Heteroassoziation bei $p = q, m = n$ und für Autoassoziation folgt aus Ansatz (4.23):

$$p = u \ln[n]/n \tag{4.26}$$

und aus (4.21)

$$M^* = -\ln[p_0] \left(\frac{n}{u \ln[n]} \right)^2. \tag{4.27}$$

Auch für endliche Systemgröße werden die optimalen Parameter durch diese Beziehungen recht genau beschrieben.

Hifi-Korrekturkapazität
Mit verrauschten Inhaltsmustern (4.18) wird die Hifi Bedingung zu

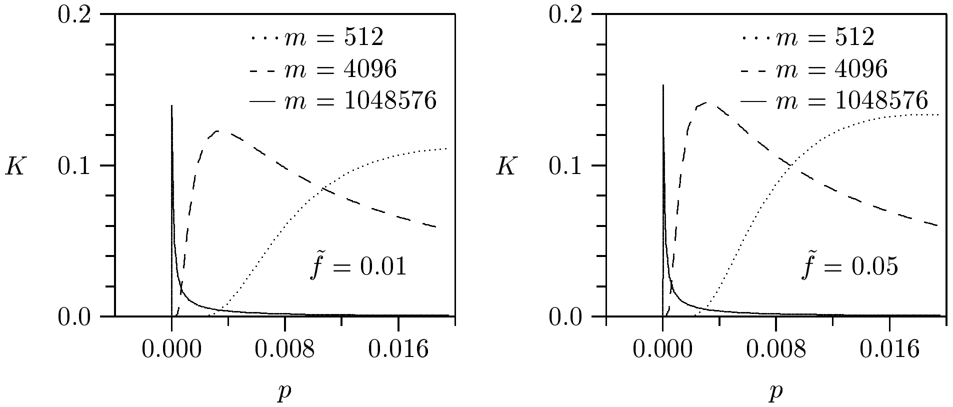

Abbildung 4.3: Korrekturkapazität in Bit/Synapse für zwei verschiedene Werte des Fehlerverhältnisses $\tilde{f}$ (links und rechts), mit der Adressierung $\hat{\beta} = 0.5$. Binärer Speicherprozess für drei verschiedene Modellgrößen m.

$\tilde{\alpha}/q = \exp\Big[mp(1-\hat{\beta})\ln[1-p_0] - \ln[q]\Big] \rightarrow 0$. Die Konstante im Ansatz (4.23) hat dann als untere Schranke

$$u > -\Big[(1-\hat{\beta})\ln[1-p_0]\Big]^{-1} \tag{4.28}$$

und wir erhalten die maximale Anzahl von Mustern als

$$M^*(\hat{\beta}) = (1-\hat{\beta})M^*, \tag{4.29}$$

wobei M^* der Wert für perfekte Adressierung aus (4.27) ist. Bei Heteroassoziation sinkt also die Ausgangskapazität linear mit steigenden Adressfehlern $A(\hat{\beta}) = (1-\hat{\beta})A$.

Bei Autoassoziation unter Hifi-Bedingung kann der Retrievalfehler in der Korrekturkapazität (3.14), wie schon bei der Ausgangskapazität, vernachlässigt werden und wir erhalten für $p \rightarrow 0$

$$\begin{aligned} K &= \max_{\hat{\beta}} \left\{ [M^*(\hat{\beta})/n][1-p(1-\hat{\beta})]\, i\left(\frac{p\hat{\beta}}{1-p(1-\hat{\beta})}\right) \right\} \\ &= \max_{\hat{\beta}} \left\{ \frac{\ln[p_0]\ \ln[1-p_0]\,(1-\hat{\beta})\hat{\beta}}{\ln[2]} \right\} = 0.17 \text{ Bit/Synapse} \end{aligned} \tag{4.30}$$

für $p_0 = 0.5$ und $\hat{\beta} = 0.5$.

In Abbildung 4.3 ist die Korrekturkapazität gegen p aufgetragen für drei endliche Systemgrößen bei konstanten Fehlerverhältnissen (wie in Abbildung 4.1). Die Adressierung ist $\hat{\beta} = 0.5$, d.h. genau die Hälfte der richtigen Einsen fehlen im Eingangsmuster. Bei dieser Adressierung erzielt das binäre Modell – auch für endliche Systemgröße – seine maximalen Kapazitätswerte. Die Korrekturkapazitätskurven verlaufen genau so wie die Ausgangskapazitätskurven. Für wachsende Systemgröße steigen die Werte der Maxima monoton und erreichen (für sehr große Systeme) den asymptotischen Wert von 0.17 Bit/Synapse, vgl. dazu Abbildung 6.17.

4.4.6 Literatur

Das binäre Speichermodell wurde zuerst von Steinbuch in [St61] vorgeschlagen. Die Analyse dieses Modells wurde von Willshaw et al [WBL69] und Palm [Pa80, Pa81, Pa91] durchgeführt und in [FM88] mit einer mathematisch strengeren Behandlung bestätigt. Eine ausführlichere Darstellung findet sich in [PS93].

4.5 Inkrementelle Speicherung

Für ein Musterpaar x^k, y^k läßt sich jedes synaptische Matrixelement zerlegen in eine Summe zweier Komponenten, der Signalkomponente, die aus dem Lernprozess dieses Musterpaares stammt und der Rauschkomponente, die aus dem Lernprozess der anderen Musterpaare resultiert. Entsprechend kann das dendritische Potential zerlegt werden in die Signalkomponente s, welche die Information über y_j^k enthält und Rauschkomponente N.

Mit den Gleichungen (2.1) und (2.6):

$$\begin{aligned} d_j &= \sum_i x_i^k c_{ij} = \sum_i \sum_l x_i^k R(x_i^l, y_j^l) \\ &= \sum_i \sum_l x_i^k R(x_i^l, y_j^l) + \sum_i x_i^k R(x_i^k, y_j^k) =: N + s \end{aligned}$$

Die Signalkomponente an On-, beziehungsweise Off-Neuronen beträgt

$$s_{on} := \sum_i x_i^k R(x_i^k, 1)\ ,\ s_{off} := \sum_i x_i^k R(x_i^k, a),$$

während die Rauschkomponente als Zufallsvariable behandelt wird, die sich an On- und Off-Neuronen gleich verhält. Es kann nun nach Gleichung (4.5) aus Abschnitt 4.2 das Signal-zu-Rausch-Verhältnis berechnet werden.

In welcher Weise sich bei gegebenem Signal-zu-Rausch-Verhältnis die Retrievalfehler auf die beiden Fehlertypen verteilen, kann durch die Schwellenseinstellung gesteuert werden. Um beide Wahrscheinlichkeiten unter 0.5 zu bringen, muß $Ed_{off} \leq \Theta \leq Ed_{on}$ erfüllt sein, was durch den Ansatz $\Theta = Ed_{off} + \vartheta\sigma(N)r = Ed_{on} - (1-\vartheta)\sigma(N)r$ mit $\vartheta \in [0,1]$ gewährleistet wird. Für große m kann das Rauschen N als Summe von einer großen Anzahl unabhängiger Zufallsvariablen betrachtet werden, für die der zentrale Grenzwertsatz gilt. Die Fehlerwahrscheinlichkeiten aus (4.4) und (4.3) lassen sich damit abschätzen:

$$\begin{aligned} \tilde{\beta} &\approx G\left[-E(d_{on} - \Theta)/\sigma(N)\right] = G\left[-(1-\vartheta)r\right] && (4.31) \\ \tilde{\alpha} &\approx G\left[-\vartheta r\right] && (4.32) \end{aligned}$$

mit der Gaussverteilung $G[x] := (1/\sqrt{2\pi}) \int_{-\infty}^{x} e^{-x^2/2} dx$.

Kleine Fehler können nur für großes Signal-zu-Rausch-Verhältnis erwartet werden, bei $r \to \infty$ für $n \to \infty$ kann das Verhalten der Fehlerwahrscheinlichkeiten mit Proposition 1 aus Anhang B approximiert werden durch

$$\tilde{\alpha} \propto \exp\left[-\frac{\vartheta^2 r^2}{2}\right] \text{ und } \tilde{\beta} \propto \exp\left[-\frac{(1-\vartheta)^2 r^2}{2}\right]. \qquad (4.33)$$

Für die verschiedenen Lernregeln aus Abschnitt 2.3.1 und die verschiedenen Fälle von Eingangsmuster, müssen nun Signal- und Rauschterm in Gleichung (4.6) analysiert werden.

Für das Ensemble von Retrievalprozessen berechnen wir das Signal-zu-Rausch-Verhältnis für eine gemittelte Anzahl aktivierter Eingangsfasern und setzen in (4.31) einen mittleren Schwellenwert $\Theta = E(\hat{k})$ ein. Wie in Abschnitt 4.4.1 b) dient das Ergebnis wieder als Abschätzung für den gemittelten Fehler eines Retrievalvorgangs, bei dem für jedes Eingangsmuster die Schwelle optimal bezüglich der Anzahl aktiver Neurone $\hat{k}$ gesetzt wird. Dieses Vorgehen ist äquivalent zum Vertauschen der Erwarungswerte bezüglich der Lernmustermittelung und der Mittelung über die Eingangsmuster in der Berechnung.

4.5.1 Berechnung des Signal-zu-Rausch-Verhältnis

Es werden wieder nacheinander die drei Adressierungsfälle aus Abschnitt 4.3 behandelt.

a) Bei dem fehlerfreien Adressmuster x^u am Eingang ist das Signal scharf determiniert zu

$$s_{on} - s_{off} = k(r_4 - r_3) - (m-k)a(r_2 - r_1),$$

r_i wie in (2.5) defininiert. Der Rauschterm zerfällt in eine Summe von unabhängigen Beiträgen aus dem Lernprozess jedes Musterpaares (x^l, y^l) mit $l \neq u$. Für jedes Paar erzeugt das Eingangsmuster am Neuron j eine Summe von k Zufallsvariablen $R(x,y)$ und von $(m-k)$ Zufallsvariablen $aR(x,y)$. Die Größe $R(x,y) = R(x_i^l, y_j^l)$ ist eine diskrete vierwertige Zufallsvariable (2.5) mit der Wahrscheinlichkeitsverteilung: $(1-p)(1-q), p(1-q), (1-p)q, pq$.

Bezeichnen wir Erwartungswert und Varianz mit $E(R)$ bzw. $\sigma^2(R)$ von $R(x,y)$ dann erhält man durch eine einfache (aber für $\sigma^2(N)$ etwas längliche) Rechnung

$$\begin{aligned} E(N) &= (M-1)\,[k + (m-k)a]\,E(R) && (4.34)\\ \sigma^2(N) &= (M-1)\left\{Q_1\sigma^2(R) + Q_2 \operatorname{Cov}[R_iR_h]\right\} && (4.35)\end{aligned}$$

mit den Abkürzungen

$$\begin{aligned} Q_1 &:= k + (m-k)a^2 \\ Q_2 &:= k(k-1) + 2ak(m-k) + a^2(m-k)(m-k-1) \\ \operatorname{Cov}[R_iR_h] &= q(1-q)\,[p(r_4 - r_3) + (1-p)(r_2 - r_1)]^2 && (4.36)\end{aligned}$$

Der Kovarianzterm: $\operatorname{Cov}[R_iR_h] := \operatorname{Cov}\Big[R(x_i^l, y_j^l)R(x_h^l, y_j^l)\Big]$ stellt ein Maß der Abhängigkeit zwischen zwei verschiedenen Beiträgen während des Lernvorgangs an den Stellen i und h in einer Spalte j der synaptischen Matrix dar.

b) Bei der Mittelung über fehlerfreie Adressmuster benutzen wir für große m wieder die Nährungen $k/m \approx (k-1)/m \approx (k+1)/m \approx p$ und $(M-1)/m \approx M/m$. Wir erhalten

$$\begin{aligned} E(s_{on} - s_{off}) &= m\,[p(r_4 - r_3) - (1-p)a(r_2 - r_1)] && (4.37)\\ E(N) &= (M-1)m\mu E(R) \end{aligned}$$

Ferner ist in Gleichung (4.35) jetzt einzusetzen:

$$Q_1 = m\left[p + (1-p)a^2\right] \quad , \quad Q_2 = m^2\mu^2 \qquad (4.38)$$

c) Für das Eingangsensemble verrauschter Adressmuster ist schließlich

$$E(s_{on}(\hat\beta) - s_{off}(\hat\beta)) = m\left[p(1-\hat\beta(1-a))(r_4 - r_3) - (1-p)a(r_2 - r_1)\right]. \qquad (4.39)$$

Bei der Beschreibung des Rauschens muß in (4.37) und (4.38) lediglich p durch $p(1-\hat\beta)$ und μ durch $\hat\mu$ ersetzt werden.

4.5.2 Signal-zu-Rausch-Verhältnis für spezielle Lernregeln

Bei den Gleichungen (4.35) und (4.37) fällt auf, daß das Signal-zu-Rausch-Verhältnis gleich bleibt wenn wir eine Lernregel R durch $bR + c$ ersetzen, wobei c eine beliebige Zahl und b eine beliebige positive Zahl sein kann. Wir nennen deshalb zwei Regeln die in einer solchen Beziehung stehen *essentiell identisch* und identifizieren sie in unseren weiteren Betrachtungen indem wir Regeln der Form

$$R = (0, r_2, r_3, r_4) \tag{4.40}$$

untersuchen. Im folgenden erweist es sich schreibetechnisch als ökonomisch, die Größen r_2, r_3, r_4 auf den abhängigen Parametersatz

$$\gamma := r_4 - r_3 - r_2 \ , \ \ \kappa := r_2 + \gamma p \ , \ \ \eta := r_3 + \gamma q$$

zu transformieren. Als Varianz einer Lernregel erhalten wir damit

$$\begin{aligned} \sigma^2(R) : &= E(R^2) - (E(R))^2 \\ &= \eta^2 p(1-p) + \kappa^2 q(1-q) + \gamma^2 p(1-p)q(1-q) \end{aligned}$$

In der Beschreibung des Eingangsensembles transformieren wir von p, a auf die Größen p, μ, siehe (4.9).

Das Signal-zu-Rausch-Verhältnis (4.6) im Fall b) schreibt sich dann als

$$r^2 = (m/M) \frac{[\mu\kappa + (1-\mu)p\gamma]^2}{[p + (\mu - p)^2/(1-p)]\,\sigma^2(R) + mq(1-q)\mu^2\kappa^2}. \tag{4.41}$$

Für Fall c) erhalten wir $r^2(\hat{\beta})$ aus Gleichung (4.41), wenn wir darin μ durch $\hat{\mu}$ (4.8) und p durch $\hat{p}$ ersetzen.

4.5.3 Optimale Lernregel

In Gleichung (4.41) kann das Signal-zu-Rausch-Verhältnis bezüglich der Parameters γ, κ und η optimiert werden, wodurch wir die *optimale Lernregel* R_0 erhalten.

Wir minimieren zunächst $\sigma^2(R)$ im Nenner des Ausdrucks für r^2 in (4.41), indem der Parameter $\eta = 0$ gesetzt wird:

$$r^2 = (m/M) \frac{[\mu\kappa + (1-\mu)p\gamma]^2}{q(1-q)\left\{[p + (\mu - p)^2/(1-p)]\,[\kappa^2 + \gamma^2 p(1-p)] + m\mu^2\kappa^2\right\}} \tag{4.42}$$

Der (große) Faktor m im zweiten Term des Nenners in Gleichung (4.42) hält das Signal-zu-Rausch-Verhältnis klein, solange nicht einer der anderen Faktoren κ oder μ verschwindet.

Auf den ersten Blick müssen wir bezüglich der mittleren Eingangsaktivität zwei unterschiedliche Fälle unterscheiden:

- Fall $\mu \neq 0$: Hier maximiert eine Regel mit $\kappa = 0$ das Signal-zu-Rausch-Verhältnis.
- Fall $\mu \to 0$: Die Eingangsaktivität verschwindet schnell genug, daß der zweite Term im Nenner vernachlässigbar bleibt. Auch in diesem Fall, wenn wir $\mu = 0$ in (4.42) einsetzen, maximiert wieder $\kappa = 0$ das Signal-zu-Rausch-Verhältnis.

Somit führen beide Fälle auf $\kappa = 0$ and $\eta = 0$ und die universelle optimale Lernregel lautet

$$R_0 = (pq, -p(1-q), -q(1-p), (1-p)(1-q)). \tag{4.43}$$

Sie wird auch als die *Kovarianzregel* bezeichnet, da bei ihr der Kovarianzterm (4.36) verschwindet.

Die *Bedingung verschwindender mittlerer Eingangsaktivität* $\mu = 0$ für bringt für alle Regeln das beste Signal-zu-Rausch-Verhältnis. Sie ist äquivalent zu $a = -p/(1-p)$ oder $p = -a/(1-a)$ und legt die optimale Kombination zwischen Eingangsaktivität und dem Modellparameter a fest. Für $p = 0.5$ erzwingt sie den Wert $a = -1$ und für $p \to 0$ den Wert $a \to 0$.

R_0 ist essentiell identisch zur Produktregel $R_0 = (1-p)(1-q)x_i y_j$, wenn die im Lernprozess angelegten Muster der Bedingung $\mu = 0$ genügen.

Für beliebiges p und a in den Inhaltsmustern, d. h. für beliebiges μ, wird das Signal-zu-Rausch-Verhältnis für die optimale Regel durch Einsetzen von R_0 in Gleichung (4.42) berechnet

$$r_0^2 = (m/M)\frac{(1-\mu)^2 p}{q(1-q)\left[p + (\mu-p)^2/(1-p)\right](1-p)}. \tag{4.44}$$

Transformiert man in (4.44) von μ nach a zurück, so erhält man

$$r_0^2 = (m/M)\frac{p(1-p)(1-a)^2}{\left[p + (1-p)a^2\right] q(1-q)}. \tag{4.45}$$

Setzt man noch zusätzlich die Bedingung $\mu = 0$ in Gleichung (4.44) ein, so erhält man schließlich das bezüglich Lernregel und Eingangsaktivität maximierte Signal-zu-Rausch-Verhältnis zu

$$r_0^2 \approx \frac{m}{Mq(1-q)}. \tag{4.46}$$

Verrauschte Adressierung
Die Optimierung des Signal-zu-Rausch-Verhältnis für den Adressierungsfall c) aus Abschnitt 4.3.1 führt wieder auf die selbe optimale Lernregel (4.43). Der Wert des Signal-zu-Rausch-Verhältnis für perfekte Adressierung wird dann durch das Adressrauschen reduziert. Mit optimaler Lernregel R_0 und $\mu = 0$ beträgt er

$$[r_0(\hat{\beta})]^2 \approx \frac{(1-p)(1-\hat{\beta})^2}{(1-\hat{\beta})(1-2p)+p} r_0^2. \tag{4.47}$$

Für alle Lernregeln mit nichtverschwindender Kovarianz (4.36), kann nur $\mu = 0$ den m^2 Term in der Varianz des Noise unterdrücken. Ist $\kappa \neq 0$ und $\mu \neq 0$, verschwindet r für $m \to \infty$. Eine weitere kurze Rechnung zeigt, daß für Lernregeln mit $\mu \neq 0$ und nichtverschwindendem γ r ebenso asymptotisch Null wird. Zusammengefaßt: nichtverschwindendes Signal-zu-Rausch-Verhältnis r ist für suboptimale Regeln nur zu erreichen falls $\mu = 0$.

4.5.4 Hebbregel und Spinregel

Offensichtlich weichen beide Regeln von der optimalen Regel R_0 in (4.43) ab. Allerdings stimmt für den Fall $p = q = 0.5$ die Spinregel mit der optimale Regel überein, es ist $R_0 =$

	Optimale Regel R_0	Hebbregel H	Spinregel C
$r^2 =$	$\frac{m}{Mq(1-q)}$	$\frac{m(1-p)}{Mq(1-pq)}$	$\frac{8mp(1-p)}{M\left[p(1-q)+(1-p)q\right]}$

Tabelle 4.1: Quadriertes Signal-zu-Rausch-Verhältnis $r^2(m, M, p, q)$ für $\mu = 0$.

$(0.25, -0.25, -0.25, 0.25)$ und für $p, q \to 0$ approximiert die Hebbregel die optimale Regel: $R_0 \to H$.

Mit Gleichung (4.41) kann auch das Signal-zu-Rausch-Verhältnis für die Hebbregel und Spinregel berechnet werden, die Ergebnisse sind in Tabelle 4.1 zusammengefasst. Wie erwartet wird die Hebbregel essentiell gleich zu R_0 für $p, q \to 0$. Im $a = 0$ Modell, in dem Parameter a nicht so adjustiert werden kann um $\mu = 0$ zu erreichen, ist die noch stärkere Forderung and die Spärlichkeit der Adressmuster $mp^2 \to 0$ notwendig. Hierdurch ist $\mu \to 0$ schnell genug erfüllt um die essentielle Gleichheit zwischen H and R_0 zu gewährleisten.

Durch Vergleich der r^2 Werte der unterschiedlichen Regeln aus Tabelle 4.1 kann die Analyse von Hebbregel und Spinregel aus den Ergebnissen der optimale Regel in Abschnitt 4.5.3 abgeleitet werden.

4.5.5 Signal-zu-Rausch-Verhältnis und Lernregeln

Mit inkrementellem Lernverfahren ergeben sich aus der Analyse des Signal-zu-Rausch-Verhältnis für Einschrittretrieval die Folgerungen:

- Erzielt eine Regel R das Signal-zu-Rausch-Verhältnis r, dann erreicht eine Regel der Form $bR + c$, mit positiver Konstante b genau den gleichen Wert. Regeln mit dem selben Signal-zu-Rausch-Verhältnis nennen wir *essentiell gleich.*
- Die Bedingung verschwindender mittlerer Eingangsaktivität $p + (1-p)a = \mu = 0$ legt unabhängig von der Lernregel die beste Kombination der Parameter p and a fest.
- Das maximale Signal-zu-Rausch-Verhältnis r_0 wird immer von der Kovarianz Regel R_0 (4.43) erreicht, sie heißt deshalb *optimale Lernregel.* Für steigendes μ fällt r_0 stetig bis $r_0 = 0$.
- Jede von R_0 verschiedene Regel liefert $r = 0$, sobald die Bedingung $\mu = 0$ verletzt ist.
- Die Hebbregel wird essentiell gleich mit R_0 für spärliche Adress- und Inhaltsmuster d.h. $q \to 0$ und $p \to 0$.
- Die Spinregel ist optimal für $p = q = 0.5$.
- *Hifi-Retrieval bei Speicherung extensiv vieler Muster*, d.h. $r \to \infty$ für $m \to \infty$ und $M/m > 0$ erreichen nur die Lernregeln R_0 und H und zwar für spärliche Inhaltsmuster $q \to 0$. Die Spinregel S erreicht höchstens $r = \text{const}$, also immer endliche Retrievalfehler für $m \to \infty$.

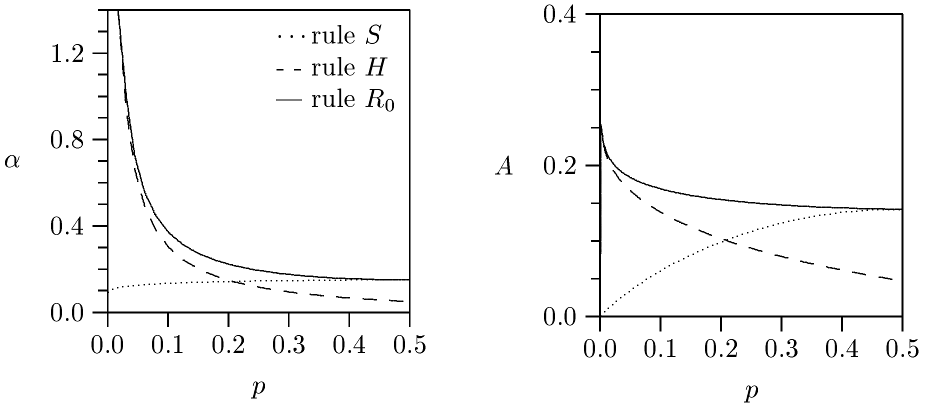

Abbildung 4.4: Maximale Anzahl speicherbarer Muster M^* (links) und Ausgangskapazität A in Bit/Synapse (rechts) für ein vorgegebenes Fehlerverhältnis $\tilde{f} = 0.01$. Inkrementeller Speicherprozess $m, n \to \infty$, Adressmuster und Inhaltsmuster mit $p = q$ und $n = m$. Die Performanz der optimalen Regel wird für $p = 0.5$ von der Spinregel und für $p \to 0$ von der Hebbregel erreicht. Was man in diesem Diagramm nicht sieht: Optimale Regel und Hebbregel erreichen bei $p \to 0$ den Wert von $A = 0.72$ Bit/Synapse.

4.5.6 Heteroassoziation

Mit perfekter Adressierung und $\mu = 0$ ergibt sich für die optimale Regel R_0 bei einem vorgegebenen Wert des Signal-zu-Rausch-Verhältnis die maximale Anzahl speicherbarer Muster aus Gleichung (4.46):

$$M^* = m/(r^2 q(1-q)). \tag{4.48}$$

Verwendet man eine Schwellenwahl so daß $\tilde{\alpha}/q = \tilde{\beta}/(1-q) =: \tilde{f}$, dann kann die Ausgangskapazität für kleines Fehlerverhältnis $\tilde{f}$ mit (3.13) und (4.48) berechnet werden

$$A \approx \frac{i(q) + q(1-q)\tilde{f}\left\{\log_2\left[q\tilde{f}\right] + \log_2\left[(1-q)\tilde{f}\right]\right\}}{r^2 q(1-q)}. \tag{4.49}$$

Ersetzt man gemäß (4.33) $r = G^{-1}\left[q\tilde{f}\right] + G^{-1}\left[(1-q)\tilde{f}\right]$ in (4.49) so erhält man für gegebenes Qualitätskriterium $\tilde{f}$ die Ausgangskapazität der optimalen Regel. ($G^{-1}[x]$ bezeichnet die inverse Gaussverteilung.) Werte für Hebbregel und Spinregel erhalten wir aus einem Vergleich der r Werte in Tabelle 4.1 in Abschnitt 4.5.4. In Abbildung 4.4 sind Ausgangskapazitäten für optimale, Hebb- und Spinregel gegen p aufgetragen.

Hifi-Ausgangskapazität

Wie bereits in Abschnitt 4.5.5 konstatiert, erfüllt bei Speicherung extensiv vieler Muster optimale Regel und Hebbregel die Vorraussetzung für Hifi-Retrieval[3] für spärliche Inhaltsmuster mit $q \to 0$. In diesem Abschnitt werden wir spärliche Inhaltsmuster vorraussetzen.

Wir machen nun für das Signal-zu-Rausch-Verhältnis den Ansatz

$$r = \sqrt{-2\ln[q]}/\vartheta, \tag{4.50}$$

---

[3] Mit der Spinregel ist ein divergierendes Signal-zu-Rausch-Verhältnis nur bei $M^*/m \to 0$, also verschwindender Ausgangskapazität zu erfüllen, diese Regel wird daher nicht weiter betrachtet.

wobei die Schwelle asymmetrisch gesetzt wird $\vartheta \to 1$, um (für spärliche Muster) ähnliche Fehlerverhältnisse an Off- und On-Neuronen zu erreichen. Dieser Ansatz ist gerade so gewählt, daß er die für Hifi-Qualität notwendige Bedingung (3.18) erfüllt[4]. Auch das Hifi-Qualitätskriterium selbst (3.17) ist erfüllt, denn $\tilde{f} \to 0$ verschwindet so schnell, daß der zweite Summand in Gleichung (4.49), der Fehlerkorrekturterm verschwindet.

Durch Einsetzen von Ansatz (4.50) in (4.48) folgt

$$M^* = m/\left(-2q(1-q)\ln[q]\right). \tag{4.51}$$

Mit (4.51) und der Definition der Musterkapazität (3.12) erhalten wir schließlich für die optimale Regel R_0 und $\mu = 0$ die asymptotische Hifi-Ausgangskapazität $A = 0.72$ Bit/Synapse. Wie man in Abbildung 4.4 sieht, wird der asymptotische Wert erst für extrem spärliche Lernmuster angenähert. Das der asymptotische Wert auch tatsächlich angenommen wird, erkennt man sogar erst auf einem Plot mit doppelt logarithmisch aufgetragenem Parameter p, vgl. [Pa92].

Bemerkung
Anders als beim binären Lernverfahren, bei dem auch die Adressmuster spärlich sein mußten, wird hier der asymptotische Kapazitätswert für Adressmuster mit beliebigem p erreicht.

Genauso wie beim binären Lernverfahren erreicht man für den wegen der Abwesenheit von Fehlertoleranz uninteressanten Fall singulärer Adressmuster und nichtspärliche Inhaltsmuster (mit beliebigen $q = \text{const}$) eine Ausgangskapazität von $A = i(q) \leq 1$ Bit/Synapse. Es erzeugen dann beide Lernvorgänge die gleiche Speichermatrix.

Fehlertoleranz
Für Heteroassoziation mit verrauschten Adressen erhalten wir die Ausgangskapazität der optimalen Regel R_0 bei $\mu = 0$ mit Gleichung (4.47) (es sei daran erinnert das $r^2 \propto m/M$):

$$A(\hat{\beta}) = \frac{(1-p)(1-\hat{\beta})^2}{p + (1-\hat{\beta})(1-2p)} A. \tag{4.52}$$

Für nichtspärliche Adressmuster mit $p = 0.5$ erhält man $A(\hat{\beta}) = (1-\hat{\beta})^2 A$ und für spärliche Adressmuster mit $p \to 0$, genau wie beim binären Lernverfahren $A(\hat{\beta}) = (1-\hat{\beta})A$.

4.5.7 Autoassoziation

Bei Autoassoziation unter dem Hifi-Qualitätskriterium erhält man analog zur Vorgehensweise in Abschnitt 4.4.5 beim Rechengang zu (4.30):

$$K(n) = \max_{\hat{\beta}} \frac{\vartheta^2(1-\hat{\beta})\hat{\beta}\log_2\left[p\hat{\beta}\right]}{2\ln[p]} \approx \max_{\hat{\beta}} \frac{\vartheta^2(1-\hat{\beta})\hat{\beta}}{2\ln[2]} = 0.18 \text{ Bit/Synapse} \tag{4.53}$$

Das Maximum wird wieder angenommen für $\hat{\beta} = 0.5$ und für $\vartheta \to 1$.

Die gleiche Optimierung in $\hat{\beta}$ kann für feste Werte von p und vorgegebenes (nichtverschwindendes) Fehlerverhältnis $\tilde{f}$ ausgeführt werden, das Ergebnis ist in Abbildung 4.5 dargestellt. Anders als beim binären Speicherprozess wird das Optimum hier für Werte von $\hat{\beta}$ angenommen, die über 0.5 liegen.

---

[4] Es folgt $q = \exp\left[-(\vartheta r)^2/2\right]$ und mit Anhang B: $\tilde{f} = \tilde{\alpha}/q \approx (\pi r^2/2)^{-1/2} \to 0$. Wenn ϑ langsam genug gegen Eins geht, sodaß $(1-\vartheta)r \to \infty$, so haben wir auch $\tilde{\beta} \to 0$.

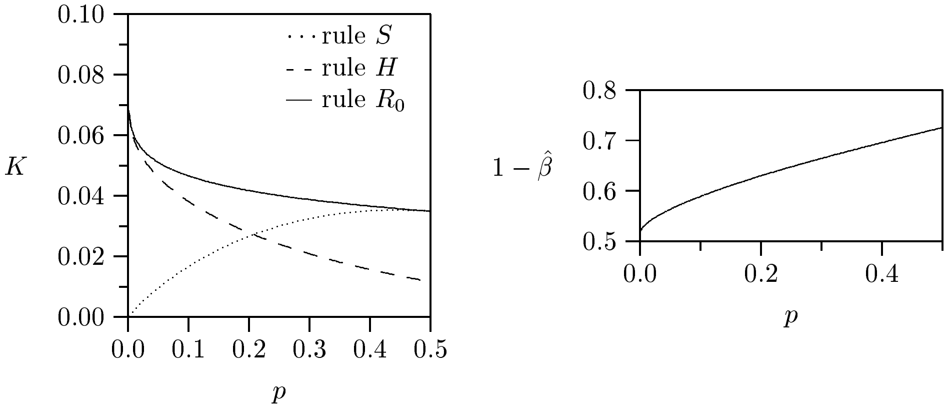

Abbildung 4.5: Korrekturkapazität in Bit/Synapse für vorgegebenes Fehlerverhältnis $\tilde{f} = 0.01$ und optimale Adressierung. Der optimale Wert für $\hat{\beta}$ in der Adressierung wurde numerisch bestimmt und ist im rechten Diagramm aufgetragen. Modell: Inkrementeller Speicherprozess, $m, n \to \infty$, Inhaltsmuster mit $q \to 0$.

	nichtspärl. Inhaltsm.	spärliche Inhaltsm.	singuläre Inhaltsm.
nichtspärl. Adressm.	—	ink. R_0	—
spärliche Adressm.	—	ink. R_0, H bin. H	—
singuläre Adressm.	ink. R_0, H bin. H	—	—

Tabelle 4.2: Modelle, die Hifi-Ausgangskapazität $A > 0$ liefern, für verschiedene Konstellationen von Adressmuster und Inhaltsmuster. Bezeichnungen: ink. = inkrementeller Speicherprozess, bin. = binärer Speicherprozess. (Beispiel: „ink.R_0, H“ bezeichnet den inkrementellen Speicherprozess, entweder mit optimaler Regel oder mit Hebbregel.)

4.6 Diskussion

4.6.1 Heteroassoziation

In Tabelle 4.2 sind für die unterschiedlichen Konstellationen von Adress- und Inhaltsmuster diejenigen Modelle eingetragen, welche einen nichtverschwindenden Wert der Hifi-Ausgangskapazität, d.h. asymptotisch $A > 0$ erreichen. Wir diskutieren die einzelnen Spalten aus Tabelle 4.2.

Nichtspärliche Inhaltsmuster
Mit singulären Adressmustern erreichen alle drei Regeln $A = 1$, aber natürlich (wegen der Adressmuster) keine Fehlertoleranz, d.h.: $A(\hat{\beta}) = 0$ für $\hat{\beta} > 0$. Bei dieser Konstellation wird jedes der M Inhaltsmuster in eine Reihe der Speichermatrix geschrieben, es findet keine verteilte Speicherung statt, binäre und inkrementelle Speicherung führen zur gleichen Synapsenmatrix und es gilt trivialerweise $A = i(q)$.

Spärliche Inhaltsmuster

	binär	inkrementell	
	H	H	R_0
nichtspärl.	—	—	$A = 0.72$
Adresse	—	—	$(1-\hat{\beta})^2$
spärliche	$A = 0.69$	$A = 0.72$	$A = 0.72$
Adresse	$(1-\hat{\beta})$	$(1-\hat{\beta})$	$(1-\hat{\beta})$

Tabelle 4.3: Hifi-Ausgangskapazität der verschiedenen Modelle für spärliche Inhaltsmuster in Bit/Synapse. Die Fehlertoleranz in der Adressierung ist in der zweiten Zeile jeder Zelle durch den Reduktionsfaktor angegeben. (Beispiel: Die Hebbregel liefert im inkrementellen Speicherprozess $A = 0.36$ Bit/Synapse wenn in der Adressierung $\hat{\beta} = 0.5$ gewählt wurde.)

In Tabelle 4.3 kann die Hifi-Ausgangskapazität für die verschiedenen Speichermodelle verglichen werden. Für spärliche Adressmuster liefern Hebbregel und Spinregel die selben Werte, da sie für $p \to 0$ essentiell gleich sind. Es ist bemerkenswert und zunächst erstaunlich, daß die Hebbregel im binären Speicherprozess beinahe die gleiche Performanz aufweist wie im inkrementellen Speicherprozess. Schließlich werden bei binärer Speicherung durch das Abschneiden der Synapseneinträge auf die Werte 0 oder 1 Informationen über die Lernmuster gelöscht und die Darstellung binärer Synapsen ist weniger aufwendig. Man sieht jedoch leicht, daß bei beiden Varianten der Speicherung die Spärlichkeitsforderung für die Muster ein extensives Wachstum der Matrixelemente verhindert; der mittlere Beitrag aller Muster an einer Synapse ist für binäre Speicherung $E(\bar{c}) = 0.5$ und für inkrementelle Speicherung $E(c) = 0.69$. Auch bei inkrementeller Speicherung besitzen also die meisten Matrixelemente die Werte 0, 1 und 2.

- Bei nichtspärlichen Adressmuster erhält man nur mit der optimalen Regel R_0 im inkrementellen Lernvorgang endliche Werte für die Ausgangskapazität.
- Alle Fälle aus Tabelle 4.3 gewährleisten Fehlertoleranz in der Adressierung.
- Solange spärliche, nicht aber singuläre Muster verwendet werden, wird eine Synapse durch mehrere Muster während des Speicherprozesses verändert, es findet also verteilte Speicherung statt.
- Eine Lockerung im Qualitätskriterium bringt keinen weiteren Anstieg der Ausgangskapazität, die Werte für eine endliche Fehlerschranke aus Abbildung 4.1 und Abbildung 4.4 übersteigen nicht die Werte aus Tabelle 4.3. Das bedeutet, daß bei wachsender Beladung asymptotisch ein scharfer Übergang stattfindet zwischen einem Bereich mit Hifi-Retrieval und einem Bereich, in dem überhaupt kein Retrieval mehr möglich ist.

Singuläre Inhaltsmuster
führen auf asymptotisch verschwindende Ausgangskapazität, weil wegen der kombinatorischen Beschränkung $M \leq n$, die Information in den Inhaltsmustern pro Synapse verschwindet.

Literatur

Die Optimierung des Signal-zu-Rausch-Verhältnisses r für inkrementelle Speicherung, bereits von Hopfield in [Hd82] vorgeschlagen, wurde von Willshaw und Dayan [WD90] und unabhängig davon von Palm [Pa90] ausgeführt. Auch Amit et al [AGS87] benutzen bereits die Kovarianzregel.

Das Signal-zu-Rausch-Verhältnis ist ein Maß dafür, wie gut das Schwellendetektionsproblem von einem Neuron gelöst werden kann. In unserem Modell wurde die Schwelle für alle Neurone auf denselben Wert Θ gesetzt, der Wert Θ wurde jedoch bezüglich der Eingangsaktivität des individuellen Eingangsmuster optimiert. In der Analyse dieses Modells konnten wir die On- und Off-Neurone am Ausgang jeweils zusammenfassen und ein gemitteltes Signal-zu-Rausch-Verhältnis behandeln. In einer Arbeit von Willshaw and Dayan [WD91] wurde für eine S/N-Analyse für ein unterschiedliches Modell ausgeführt. In ihrem Modell konnte die Schwelle für jedes Neuron einzeln auf einen Wert Θ_j gesetzt werden, aber die Schwellenwahl nicht der gemittelten Eingangsaktivität eines individuellen Eingangsmusters angepasst werden. Dadurch wurde das Signal-zu-Rausch-Verhältnis an einem einzelnen Neuron für eine gemittelte Eingangsaktivität optimiert.

Wegen dieses Unterschieds stimmen die Ergebnisse nur für $\mu = 0$ und im thermodynamischen Limes überein; dann ist auch im Willshaw-Dayan Modell die Regel R_0 optimal und erzielt gleiches Signal-zu-Rausch-Verhältnis.

Das Willshaw-Dayan-Modell ist im allgemeinen nicht invariant gegenüber der Addition mit einer Konstanten in der Lernregel. Fluktuationen der Anzahl aktivierter Neurone in einem individuellen Eingangsmuster führen für $E(R) \neq 0$ zu einer zusätzlichen Streuung im dendritschen Potential, welche nicht durch eine angepasste Schwellenwahl kompensiert werden kann.

Viele der hier dargestellten Resultate zur Heteroassoziation in assoziativen Speichermodellen findet man in Arbeiten von [Pe88], Nadal und Toulouse [NT90], Willshaw und Dayan [WD90] und Palm [Pa90, Pa91]. Das sogenannte „error-full regime“ von Nadal and Toulouse entspricht in unserer Sprechweise einem Qualitätskriterium, daß $\tilde{\alpha}, \tilde{\beta} \to 0.5$ zuläßt. Hiermit erhält man eine Ausgangskapazität, die mit $A = 0.46$ Bit/Synapse, siehe [Pa91], unter dem Hifi-Wert liegt. Diesem Fall kommt also außer der Tatsache, daß er ohnehin für Anwendungen uninteressant ist, auch im Hinblick auf das theoretische Maximum der Speicherkapazität keine Bedeutung zu.

Für binäre Speicherung mit der Hebbregel wurde der Wert der Hifi-Ausgangskapazität bereits in den alten Arbeiten von Willshaw et al [WBL69] berechnet.

4.6.2 Autoassoziation

Endliche Werte der asymptotischen Hifi-Korrekturkapazität werden bei Einschrittretrieval für spärliche Muster erreicht, die Hebbregel stellt sich für solche Inhaltsmuster als die optimale Lernregel heraus. Asymptotisch erreicht die Hebbregel bis zu $K = 0.18$ Bit/Synapse im inkrementellen Speicherprozess und bis zu $K = 0.17$ Bit/Synapse im binären Speicherprozess. Das Optimum wird für eine Adressierung mit $\hat{\beta} = 0.5$ erreicht; für Eingangsmuster bei denen im Retrieval die Hälfte der Einskomponenten rekonstruiert werden muß.

Autoassoziation mit spärlichen Mustern bildet den Gegenstand der folgenden Teile dieser Arbeit. Durch die hohen Kapazitätswerte von Einschrittretrieval liegt es im Hinblick auf Anwendungen nahe, komplexere Auslesemethoden, insbesondere iteratives Retrieval und das Verhalten bei endlicher Systemgröße zu untersuchen.

Kapitel 5

Fixpunktkapazität

5.1 Zusammenfassung

In diesem Kapitel wird eine Methode zur asymptotischen Berechnung der Kapazität für Autoassoziation spärlicher Lernmuster vorgestellt, die erstmals in [PS92, PS91] beschrieben wurde.

Speicherung mit der Hebbschen Lernregel wird zusammen mit einem von bisher beschriebenen Retrievalverfahren unterschiedlichen Ausleseverfahren untersucht, dem sogenannten *Fixpunkt-Verfahren.* Beim Fixpunkt-Verfahren besteht die Speicheranfrage nicht in einer einzigen Adresse pro Lernmuster, vielmehr wird durch eine sehr große Zahl von Speicheranfragen der gesamte Adressraum auf Fixpunkte der Retrievaldynamik abgesucht. Die Menge der gefundenen Fixpunkte stellt die Speicherantwort dar.

Werden nur die Inhaltsmuster selbst auf ihre Fixpunkteigenschaft geprüft, ist die Kapazität für dieses Verfahren mit der Musterkapazität für Fixpunktretrieval vergleichbar und wir erhalten auch die selben Zahlenwerte [WBL69, TF88, Hr89].

Die Korrekturkapazität für den Fall, daß mit dieser Methode die Fixpunkte aus dem gesamten Raum spärlicher Muster extrahiert werden nennen wir die Fixpunktkapazität F. Für das binäre Lernverfahren erhalten wir $F = 0.35$ Bit/Synapse und für das inkrementelle $F = 0.36$ Bit/Synapse. Dies sind genau die doppelten Werte, die man für Einschrittretrieval erhält. Die Fixpunktkapazität kann als obere Schranke für die Korrekturkapazität angesehen werden, die mit allen solchen Retrievalverfahren erreicht werden kann, die Fixpunkte als Ausgangsmuster liefern. Die erhaltenen Werte der Fixpunktkapazität zeigen, daß mit solchen Ausleseverfahren höhere Kapazitätswerte erreichbar sind als durch Einschrittretrieval. Es wird dann Gegenstand von Kaptitel 6 sein, zu überprüfen, ob auch bis zum Fixpunkt geführtes iteratives Retrieval zu diesen Verfahren zählt.

5.2 Fixpunkt-Ausleseverfahren

Wir betrachten wieder das Speichermodell, nachdem Inhaltsmuster $\mathcal{S} = \{x^1, ..., x^M\}$ mit konstanter Aktivierung k, d.h.: $\mathcal{S} \subset \mathcal{M}_k := \{x \in \{a,1\}^n : |x| = k\}$ mit der Hebbregel entweder durch binäre oder inkrementelle Speicherung eingespeichert wurden.

Das *Fixpunkt-Ausleseverfahren* (Abbildung 5.1) bestimmt in einer Menge von Testadressierungsmuster $\hat{\mathcal{S}}$ die darin enthaltenen Fixpunkte des Systems, indem jede Testadresse auf die Fixpunkteigenschaft geprüft wird. Ein Eingangsmuster ist Fixpunkt, wenn das im Retrieval hervorgerufene Ausgangsmuster wieder fehlerfrei mit ihm übereinstimmt. Die Ausgangsmuster dieser Auslesemethode ist also durch die Menge $\tilde{\mathcal{S}} := \hat{\mathcal{S}} \cap \mathcal{F}$ gegeben, wobei

Abbildung 5.1: Fixpunkt-Ausleseverfahren: $\mathcal{S}$ ist die Menge der Inhaltsmuster, mit der im Lernprozess die synaptische Kopplungsmatrix erzeugt wurde. Beim Fixpunkt-Ausleseverfahren werden als Ausgangsmuster alle Muster aus der Menge $\hat{\mathcal{S}}$ extrahiert, welche die Fixpunkteigenschaft erfüllen. Die Menge der Ausgangsmuster ist also die Schnittmenge aus Adressmuster und Fixpunkten $\tilde{\mathcal{S}} = \hat{\mathcal{S}} \cap \mathcal{F}$. Die Ausgangsmuster können unterteilt werden in solche, die mit den Lernmustern übereinstimmen, die Retrievalerfolge und den Rest, die sogenannten 'spurious states'.

$\mathcal{F} := \{\hat{z} \in \mathcal{M} : \tilde{z} = \hat{z}\}$ die Menge sämtlicher Fixpunkte des Systems ist. Auch als Eingangsmuster sind nur Muster der Aktivierung k zugelassen, also $\hat{\mathcal{S}} \subset \mathcal{M}_k$.

Wieder sind zwei Arten von Auslesefehlern möglich, die aber hier nicht eine einzelne Musterkomponente, sondern die Erkennung eines ganzen Musters betreffen. Einerseits können Inhaltsmuster die Fixpunkteigenschaft nicht erfüllen und andererseits können Fixpunkte auftauchen, die nicht mit einem Inhaltsmuster identisch sind, sie werden in der physikalischen Literatur oft als *spurious states* bezeichnet. Das Fixpunkt-Ausleseverfahren läßt sich durch einen binären Informationskanal mit den Fehlerwahrscheinlichkeiten

$$p_{um} = p[z \notin \tilde{\mathcal{S}} \,|\, z \in \mathcal{S}] \quad und \quad p_{ss} = p[z \in \tilde{\mathcal{S}} \,|\, z \notin \mathcal{S}] \tag{5.1}$$

beschreiben.

5.3 Informationskanal

Analog zu Abschnitt 3.3 modellieren wir den Datenfluß durch einen Informationskanal mit binären Kanalkomponenten. Aber anders als dort überträgt der Informationskanal nicht ein einzelnes Muster, sondern die gesamte Mustermenge $\mathcal{S}$. Hierzu wird die komplette Mustermenge $\mathcal{S}$ in einen (sehr hochdimensionalen) binären Vektor X kodiert, indem für jedes Element i, das zu $\mathcal{S}$ gehört $X_i = 1$ gesetzt wird und sonst $X_i = 0$. Die Dimension von X ist $\binom{n}{k}$, sie ist gleich der Anzahl der Elemente in der Menge $\mathcal{M}_k$, der Menge aller potentiell zu $\mathcal{S}$ gehörigen Muster.

Durch den Informationskanal, der aus mindestens $\sharp\mathcal{M}_k$ binären Kanalkomponenten bestehen muß, wird Vektor X in gleicher Weise übertragen, wie ein Einzelmuster in Abschnitt 3.3, vgl. Abbildung 3.3. Jedes nicht fehlerfrei mit einem Inhaltsmuster übereinstimmende Muster in der Ausgangsmenge wird pauschal als Retrievalfehler gewertet, ganz unabhängig vom Abstand zu seinem nächsten Inhaltsmuster.

Die bedingte Information erhalten wir analog zu Gleichung (3.7), nun mit den Fehlerwahrscheinlichkeiten (5.1)

$$I\left(\mathcal{S} \mid \tilde{\mathcal{S}}\right) = \sharp\mathcal{M}\, i(\, p[\mathcal{S}], p_{um}, p_{ss}) =: I_{um} + I_{ss}. \tag{5.2}$$

Die Terme I_{um} und I_{ss} korrespondieren zu den beiden unterschiedlichen Fehlertypen, den instabilen Lernmustern $\mathcal{S} \setminus \tilde{\mathcal{S}}$ und den spurious states $\tilde{\mathcal{S}} \setminus \mathcal{S}$. Die Ausdrücke dafür erhalten wir einfach durch Abschätzung der Kardinalitäten von den Mengen in Abbildung 5.1:

$$\begin{aligned} I_{um} &= \sharp\left\{\mathcal{M}_k \setminus \tilde{\mathcal{S}}\right\} i\left(\frac{\sharp\left\{\mathcal{S} \setminus \tilde{\mathcal{S}}\right\}}{\sharp\left\{\mathcal{M}_k \setminus \tilde{\mathcal{S}}\right\}}\right) && (5.3)\\ I_{ss} &= \sharp\tilde{\mathcal{S}}\; i\left(\frac{\sharp\left\{\tilde{\mathcal{S}} \setminus \mathcal{S}\right\}}{\sharp\tilde{\mathcal{S}}}\right). && (5.4)\end{aligned}$$

5.4 Hifi-Qualitätskriterium

Wir untersuchen das Fixpunkt-Ausleseverfahren Hifi-Kapazitätswerte für die beiden Extremmöglichkeiten der Adressierung: perfekt, also fehlerfrei und vollständig unscharf. Für diese Fälle werden Bedingungen an die Fehlerwahrscheinlichkeiten (5.1) abgeleitet, welche die Hifi-Bedingung (3.17) erzwingen.

5.4.1 Perfekte Adressierung

Für perfekte Adressierung, d.h. $\hat{\mathcal{S}} = \mathcal{S}$, gilt für die Fehlerwahrscheinlichkeit $p_{ss} = 0$, da wegen $\tilde{\mathcal{S}} = \mathcal{S} \cap \mathcal{F}$ keine spurious states als Ausgangsmuster auftreten können, vgl. Abbildung 5.2.

Die bedingte Information dann allein durch Gleichung (5.3) gegeben

$$I_{um}\left(\mathcal{S} \mid \tilde{\mathcal{S}}\right) = \binom{n}{k} p_{um}\; p[\tilde{\mathcal{S}}]\; \mathrm{ld}\left[\frac{p[\bar{\tilde{\mathcal{S}}}]}{p_{um}\; p[\mathcal{S}]}\right]. \tag{5.5}$$

Wenn wir ausnützen, daß $\sharp\mathcal{S}, \sharp\tilde{\mathcal{S}} << \sharp\mathcal{M}_k$, so erhalten wir

$$I_{um}/n^2 = M p_{um}\; ld\left[\frac{\binom{n}{k}}{M p_{um}}\right] / n^2.$$

Die Hifi Bedingung (3.17) ist nun erfüllt, falls die Fehlerwahrscheinlichkeit p_{um} der Bedingung genügt

$$\frac{k\; ld\,[n]\, M}{n^2} p_{um} \to 0 \quad \text{für} \quad n \to \infty. \tag{5.6}$$

Abbildung 5.2: Fixpunkt-Ausleseverfahren für perfekte Adressierung: Hier ist $\hat{\mathcal{S}} = \mathcal{S}$, es werden also lediglich die Inhaltsmuster auf Fixpunkteigenschaft geprüft. Bei dieser Adressierung können prinzipiell keine spurious states auftreten, wir haben es also mit nur einem Fehlertyp zu tun, den instabilen Inhaltsmustern.

5.4.2 Unscharfe Adressierung

Für völlig unscharfe Adressierung mit $\hat{\mathcal{S}} = \mathcal{M}_k$, dargestellt in Abbildung 5.3, kann durch die Fixpunkt-Auslesemethode Information über die Gesamtmenge der Lernmuster extrahiert werden. Sie kann jedoch nicht zur Vervollständigung von Einzelmustern verwendet werden. Die Hifi-Korrekturkapazität für diesen Adressierungsfall, von uns *Fixpunktkapazität* genannt, soll nun berechnet werden. Da keine Transinformation über die Inhaltsmuster in der Adressierung steckt, ist sie gleich mit der Ausgangskapazität.

Es tragen hier beide Terme in Gleichung (5.2) für die bedingte Information bei

$$I_{ss} = \binom{n}{k}\left\{ p_{ss}\, p[\bar{\mathcal{S}}]\ \mathrm{ld}\left[\frac{p[\tilde{\mathcal{S}}]}{p_{ss}\, p[\bar{\mathcal{S}}]}\right] + \left(p[\tilde{\mathcal{S}}] - p_{ss}\, p[\bar{\mathcal{S}}]\right)\ \mathrm{ld}\left[\frac{p[\tilde{\mathcal{S}}]}{p[\tilde{\mathcal{S}}] - p_{ss}\, p[\bar{\mathcal{S}}]}\right]\right\}.$$

Mit $\sharp\mathcal{S} << \sharp\mathcal{M}_k$, der Abschätzung $p[z \notin \mathcal{S}] \approx 1$ und $p[z \in \tilde{\mathcal{S}}] \approx p[z \in \tilde{\mathcal{S}} \cap z \notin \mathcal{S}] + p[z \in \mathcal{S}]$ welche gilt, falls $p_{um} \to 0$, erhalten wir

$$I_{ss}/n^2 = \binom{n}{k}\frac{p_{ss}}{n^2}\ \mathrm{ld}\left[1 + \frac{M}{\binom{n}{k}p_{ss}}\right] + \frac{M}{n^2}\ \mathrm{ld}\left[1 + \frac{\binom{n}{k}p_{ss}}{M}\right].$$

Da für alle betrachteten Lernregeln die Ordnung von M kleiner als n^2 ist, können wir mit $v := \binom{n}{k}p_{ss}/n^2$ abschätzen:

$$I_{ss}/n^2 \leq v\ \mathrm{ld}[1 + 1/v] + \mathrm{ld}[1 + v]\,.$$

Die Majorante verschwindet nur bei $v \to 0$ für $n \to \infty$.

Abbildung 5.3: Fixpunkt-Ausleseverfahren mit völlig unscharfer Adressierung: Für den Fall $\hat{\mathcal{S}} = \mathcal{M}_k$ werden sämtliche spärlichen Muster auf Fixpunkteigenschaft geprüft.

Die Hifi-Bedingung (3.18) ist erfüllt, falls p_{um} der Bedingung (5.6) genügt, sowie p_{ss} der folgenden Bedingung:

$$v \approx \binom{n}{k} p_{ss} \to 0 \ \text{ für } \ n \to \infty. \tag{5.7}$$

Die Bedingungen (5.6) und (5.7) werden in den Abschnitten 5.6 und 5.7 für die Berechnung von Musterkapazität und Fixpunktkapazität für die Hebbregel ausgenutzt.

5.5 Auslesefehlerwahrscheinlichkeiten

Die Fehlerwahrscheinlichkeiten p_{um} und p_{ss} können einfach auf die Fehlerwahrscheinlichkeiten $\tilde{\alpha}$ und $\tilde{\beta}$ des Einschrittretrieval aus Kapitel 4 zurückgeführt werden, da der Fixpunkttest lediglich einen Retrievalschritt benötigt.

Zunächst führen wir die Wahrscheinlichkeit dafür ein, daß alle On-Neurone bzw. alle Off-Neurone eines Eingangsmusters durch einen Retrievalschritt hindurch erhalten bleiben.

Wir betrachten ein Eingangsmuster der Form $z = (1, ..., 1, a, ..., a)$, die wir ohne Beschränkung der Allgemeinheit annehmen können. Falls z kein Inhaltsmuster ist, erhält man:

$$Q_{ss}^{on} \ := \ p\left[\sum_{i=1}^{k} c_{ij} \geq \Theta \ \text{ für } \ j = 1, ..., k \mid z \notin \mathcal{S}\right] \tag{5.8}$$

$$Q_{ss}^{off} \ := \ p\left[\sum_{i=1}^{k} c_{ij} < \Theta \ \text{ für } \ j = k+1, ..., n \mid z \notin \mathcal{S}\right]. \tag{5.9}$$

Falls z Inhaltsmuster ist, so erhält man:

$$Q_{ms}^{on} := p\left[\sum_{i=1}^{k} c_{ij} \geq \Theta \text{ für } j = 1, ..., k \mid z \in \mathcal{S}\right] \tag{5.10}$$

$$Q_{ms}^{off} := p\left[\sum_{i=1}^{k} c_{ij} < \Theta \text{ für } j = k+1, ..., n \mid z \in \mathcal{S}\right]. \tag{5.11}$$

Die Fehlerwahrscheinlichkeiten des Fixpunkt-Ausleseverfahren können mit diesen Definitionen zerlegt werden:

$$p_{um} = 1 - Q_{ms}^{on} Q_{ms}^{off} \tag{5.12}$$

$$p_{ss} = Q_{ss}^{on} Q_{ss}^{off} \tag{5.13}$$

Zur Berechnung der Q-Wahrscheinlichkeiten setzen wieder asymptotische Unabhängigkeit von k synaptischen Matrixelementen c_{ij} einer Spalte voraus, eine Annahme, die für spärliche Muster gerechtfertigt ist, siehe Anhang A. Für Off-Neuronen gilt

$$Q_{ss}^{off} = Q_{ms}^{off} = (1 - \tilde{\alpha})^{n-k}, \tag{5.14}$$

weil die Schwellendetektionsprobleme der Off-Neurone (Matrixspalten $k+1, ..., n$) als statistisch unabhängig betrachtet werden können und mit Gleichung (4.4) folgt
$p[\sum_{i=1}^{k} c_{ij} z_j \geq \Theta \mid z \notin \mathcal{S}] \approx p[\sum_{i=1}^{k} c_{ij} x_j^l \geq \Theta \mid j \in \{j : x_j^l = 0\}] = \tilde{\alpha}$.

Im Gegensatz dazu sind die Schwellendetektionsprobleme der einzelnen On-Neurone (Matrixspalten $1, ..., k$) jedoch nicht statistisch unabhängig, da für lokale Lernregeln die Matrix c_{ij} symmetrisch ist. Hier kann lediglich nach unten abgeschätzt werden: $Q_{ss}^{on} \geq \tilde{\alpha}^k$ und $Q_{ms}^{on} \geq (1 - \tilde{\beta})^k$.

5.6 Binäre Speicherung

Wir greifen nun auf Abschnitt 4.4 zurück: Setzt man die Schwelle $\Theta = k$ dann sind für $p_0 = 0.5$ die Fehlerwahrscheinlichkeiten $\tilde{\alpha} = 2^{-k}$, sowie $\tilde{\beta} = 0$. Es gilt $Q_{ms}^{on} = 1$ und wegen der Symmetrie der Matrixelemente

$$Q_{ss}^{on} = 2^{-k^2/2}. \tag{5.15}$$

Um die Anzahl speicherbarer Muster wieder groß zu machen, verwenden wir auch hier spärliche Muster gemäß Ansatz (4.26): $k = u \ \ln[n]$.

Musterkapazität
Setzen wir (4.26) und (4.27) in (5.6) ein, so ist

$$\frac{\ln[2]}{k} \operatorname{ld}[n] \, p_{um} = p_{um}/u \to 0 \text{ für } n \to \infty.$$

Mit (5.14) und (5.12) erhalten wir schließlich die Bedingung an die Fehlerwahrscheinlichkeit $\tilde{\alpha}$, die für das Fixpunkt-Ausleseverfahren das Hifi-Kriterium (5.6) garantiert

$$n \, \tilde{\alpha} \to 0 \text{ für } n \to \infty. \tag{5.16}$$

Die Forderung (5.16) ist mit der Hifi-Bedingung fürs Einschrittretrieval identisch, so daß wir für das Fixpunkt-Ausleseverfahren den selben Wert der Musterkapazität (3.12) erhalten

$$P = \ln[2] = 0.69 \text{ Bit/Synapse.} \tag{5.17}$$

Fixpunktkapazität
Mit der Beziehung (5.15) können wir Gleichung (5.13) schreiben als

$$p_{ss} \approx \exp\left[-\frac{\ln[2]}{2}\left(u \ln[n]\right)^2\right].$$

Setzen wir diesen Ausdruck in die Bedingung (5.7) ein, dann erhalten wir als führenden Exponenten $u(\ln[n])^2 \left[1 - u\ln[2]/2\right]$. Damit dieser Exponent negativ wird, muß die noch stärkere Bedingung an die Konstante im Ansatz (4.26) erfüllt sein, nämlich $u > 2/\ln[2]$. Mit dieser Zwangsbedingung finden wir für die Fixpunktkapazität den Wert

$$F = \frac{\ln[2]}{2} = 0.35 \text{ Bit/Synapse.} \tag{5.18}$$

5.7 Inkrementelle Speicherung

Für die Hebbregel im inkrementellen Lernprozess sind die Fehlerwahrscheinlichkeiten für große n durch die Gleichung (4.33) gegeben.

Musterkapazität
Verwenden wir den Ansatz

$$M = b \; n^2/(k \ln[n]), \tag{5.19}$$

worin b eine positive Zahl ist, so reduziert sich die Bedingung (5.6) wieder auf die Hifi-Bedingung fürs Einschrittretrieval, wie schon im binären Fall. Setzt man (5.19) in (4.46) und (4.33), dann wird die Hifi-Bedingung erfüllt, indem man $b < \vartheta^2/2$ wählt. Auch bei inkrementeller Speicherung erhalten wir also mit (3.12) für das Fixpunkt-Ausleseverfahren den gleichen Wert der Musterkapazität wie bereits für Einschrittretrieval

$$P = \frac{1}{2\ln[2]} = 0.72 \text{ Bit/Synapse.} \tag{5.20}$$

Fixpunktkapazität
In Proposition 2 des Anhangs B wird gezeigt, daß sich die Fehlerwahrscheinlichkeit p_{ss} verhält wie

$$p_{ss} \propto \exp\left[-\frac{\vartheta^2}{4M}n^2\right] \quad \text{für } n \to \infty.$$

In der Bedingung (5.7) ergibt sich hiermit der führende Exponent zu $k \ln[n] - \frac{\vartheta^2}{4M}n^2$. Er wird negativ für $b < \vartheta^2/4$ und damit erhalten wir für die Fixpunktkapazität

$$F = \frac{1}{4\ln[2]} = 0.36 \text{ Bit/Synapse.} \tag{5.21}$$

5.8 Diskussion

Für perfekte Adressierung und im Fall des fehlerfreien Auslesens – die Wahrscheinlichkeit dafür geht durch die Hifi-Bedingung im Limes $n \to \infty$ gegen Eins – unterscheiden sich die drei Auslesemethoden Einschrittretrieval, iteratives Retrieval und das Fixpunkt-Ausleseverfahren nicht; in jedem Fall wird die Menge der Inhaltsmuster reproduziert. Es ist demnach zu erwarten, daß die Analyse jeweils die gleichen Werte der Musterkapazität liefert und die Übereinstimmung unserer Ergebnisse mit [WBL69, Pa91] bestätigt diese neue Methode der Analyse.

Bei den Werten für die Fixpunktkapazität handelt es sich um neue Resultate. Das Fixpunkt-Ausleseverfahren selbst kann für das (schnelle) Wiedererkennen gelernter Muster verwendet werden, es eignet sich jedoch nicht für Mustervervollständigung.

Durch die vollständige Explorierung der Fixpunkte des Fixpunkt-Ausleseverfahrens stellt die Fixpunktkapazität gleichzeitig eine obere Schranke an die Hifi-Korrekturkapazität dar, welche bei Mustervervollständigung mit beliebigen Retrievalverfahren erreicht werden kann, die bis zum Fixpunkt geführt werden. Die Zahlenwerte dieser Schranken, die im Vergleich zur Korrekturkapazität für Einschrittretrieval aus Kapitel 4, (4.30) und (4.53) genau doppelt so groß sind, weisen darauf hin, daß durch iterativem Retrieval höhere Kapazitätswerte zu erzielen sind. Freilich muß untersucht werden, inwieweit diese Schranken von iterativen Retrievalmethoden bei Mustervervollständigung überhaupt erreichbar sind. Wir wollen dieser Frage im nächsten Kapitel nachgehen.

Kapitel 6

Iteratives Retrieval

6.1 Fragestellung

In diesem Kapitel wird die Anzahl der Modellvarianten eingeschränkt. Wir behandeln Autoassoziation, wo iteratives Retrieval in natürlicher Weise anwendbar ist, indem das Ausgangsmuster als neues Adressmuster verwendet wird u.s.w. Es werden nur solche Modellvarianten und Parameterbereiche betrachtet, die sich bereits im Einschrittretrieval und im Fixpunkt-Verfahren als besonders effektiv erwiesen haben. Bei Hebb-Speicherung von Zufallsmustern im (a=0)-Modell erreicht man bereits mit Einschrittretrieval sehr hohe asymptotische Korrekturkapazitäts-Werte, nämlich um $K = 0.17$ Bit/Synapse, siehe Kapitel 4. Auch das Fixpunkt-Ausleseverfahren arbeitet effektiv, in Kapitel 5 wurden Werte um $F = 0.35$ Bit/Synapse ermittelt. Diese Werte werden für *spärliche Muster* mit konstanter extrem kleiner Aktivierung, d.h.: $\sum_{i=1}^{n} x_i^\nu = k \;\; \forall\, \nu$ erreicht, bei denen $k \propto \log[n]$ gilt und für ein Qualitätskriterium, welches asymtotisch verschwindende Retrievalfehler fordert.

In diesem Modell soll nun iteratives Retrieval untersucht werden. Wir schlagen dafür drei Regelungsstrategien für die Schwelle vor, die sich experimentell bewährt haben. Mit elementaren statistischen Methoden wird für den zweiten Retrievalschritt eine Analyse im asymptotischen Fall und für das endliche Systems durchgeführt. Die Ergebnisse dieser Analyse werden zusammen mit Ergebnissen aus Simulationsexperimenten anhand zahlreicher Diagramme dargestellt und diskutiert.

6.2 Schwellenstrategien

Gute Ergebnisse bei Einschrittretrieval erfordern eine Optimierung der Schwellenwahl vom Benutzer. Für Lernmuster mit konstanter Aktivierung k sollte die Schwelle so eingestellt werden, daß die Aktivierung im Ausgangsmuster mit derjenigen im Lernmuster übereinstimmt.

Bei iterativem Retrieval ist die Schwelle jedes Iterationsschrittes, sowie der Zeitpunkt des Abbruchs der Iterationsschleife von außen zu bestimmen; diese Steuerung beim iterativen Retrieval bezeichnen wir als *Schwellenstrategie*.

Strategie CA (constant activity)
Für alle Modellvarianten ist eine Schwellenstrategie durchführbar, die in jedem Retrievalschritt die Aktivierung des Ausgangsmusters möglichst nahe dem Wert der Lernmuster hält. Wie beim Einschrittretrieval wird die (diskrete) Größe $\left|\hat{k}(\Theta) - k\right|$ durch ein Austesten verschiedener Schwellenwerte minimiert.

Strategie Tk
Nach binären Lernverfahren (2.9) gilt für die dendritischen Potentiale aller Neurone $d_j \leq \hat{k}$, wobei $\hat{k}$ die Aktivierung der Adresse ist. Stimmen im ersten Schritt l Einsen mit dem Lernmuster überein, so ist es günstig, die Schwelle auf den Wert l zu setzen, womit sichergestellt ist, daß im Ausgangsmuster alle On-Neurone des Lernmusters aktiviert sind. Da beim verrauschten Anfangsmuster die Zahl l unbekannt ist, muß die Schwelle wieder durch Austesten gesucht werden. Sie wird auf den größten Integerwert gesetzt, bei dem die Ausgangssaktivierung höher liegt als die Aktivierung in den Inhaltsmustern. Damit erreicht man in den meisten Fällen die optimale Schwellenwahl. Bei sehr hoher Beladung und einem hohen Anteil falscher Einsen im Anfangsmuster kann jedoch der so gefundene Schwellenwert zu hoch sein. In darauf folgenden Retrievalschritten können dann durch ein Anheben der Schwelle auf den Wert k falsche Einsen eliminiert werden, ohne die On-Komponenten zu verlieren.

Strategie Tk+
Verbesserte Version von Strategie Tk. Es zeigte sich in Abschnitt 4.4.3, daß die Fehlertoleranz des Assoziativspeichers dann am schlechtesten ist, wenn die Adresse durch zusätzlich eingestreute Einsen verrauscht ist. Genau solche Adressen liegen ab dem zweiten Schritt in Strategie Tk vor, sofern die Schwelle im ersten Schritt richtig gesetzt wurde. Tatsächlich werden wir in der Simulation mit Strategie Tk bei hoher Beladung ab dem zweiten Schritt eine Explosion der Aktivierung im Ausgangsmuster beobachten, siehe Abschnitt 6.4.2, wodurch im iterativen Retrieval sogar schlechtere Resultate erzielt werden als mit Einschrittretrieval.

Um dies zu vermeiden, verwendet Strategie Tk+ ab dem zweiten Retrievalschritt zusätzlich eine logische UNDung zwischen Eingangsmuster und Ausgangsmuster. Hierdurch wird ein weiteres Anwachsen der Aktivierung in den Ausgangsmustern verhindert. Falls die Schwelle im ersten Schritt korrekt gesetzt wurde, daß danach alle On-Komponenten des Inhaltsmusters bereits vorhanden sind, dann sind alle in späteren Schritten neu hinzukommenden Einsen falsch und ihre Unterdrückung durch die UNDung bringt eine erhebliche Verbesserung.

6.3 Statistische Analyse des zweiten Schrittes

Die Einschrittanalyse aus Kapitel 4 kann nicht unverändert auf einen iterierten Retrievalschritt angewendet werden, wegen der Korrelationen, welche dann zwischen dem Noise im Eingangsmuster und dem synaptischen Noise (von den anderen Lernmustern) bestehen.

Wir untersuchen die ersten zwei Schritte für Retrieval mit der Strategie Tk+. Das Inhaltsmuster x sei von der Form $(1, 1, 1, ..., a, a, a)$, es besitzt wieder die Aktivierung k. In den Abbildungen 6.1 und 6.2 sind erster und zweiter Retrievalschritt schematisch dargestellt. Es werden Erwartungswerte von den Besetzungszahlen der aktiven Gebiete in den Ein- und Ausgangsmuster $\hat{x}$, $\tilde{x}$ und $\tilde{\tilde{x}}$ betrachtet. Zur besseren Übersicht sind diese Gebiete[1] zusammenhängend dargestellt. Im ersten Schritt wird ein verrauschtes Muster $\hat{x}$ mit $\hat{\alpha} = 0$ und der Aktivierung $\hat{m}_{on} = k(1-\hat{\beta})$ als Adresse[2] angelegt. Das Ausgangsmuster $\tilde{x}$ enthält $\tilde{m}_{on}$ neuaktivierte On-Bits und $\tilde{m}_{off}$ falsche Off-Bits, die Eigenschaften lassen sich gemäß Abschnitt 4.4 bestimmen, nach (4.18)

$$\tilde{\alpha} = (1 - p_0)^{(1-\hat{\beta})k}.$$

Im zweiten Schritt gilt wegen der Undungsvorschrift zwischen Ein- und Ausgangsmuster

$$\tilde{\tilde{\alpha}} = \tilde{\alpha}\, p[d_j \geq k \mid d_j \geq (1 - \hat{\beta})k]. \tag{6.1}$$

---

[1] So bezeichnet z. Bsp. $\tilde{m}_{on}$ die Anzahl der aktiven Neurone in $\tilde{x}$.

[2] Definition von $\hat{\alpha}$ und $\hat{\beta}$, siehe (4.7).

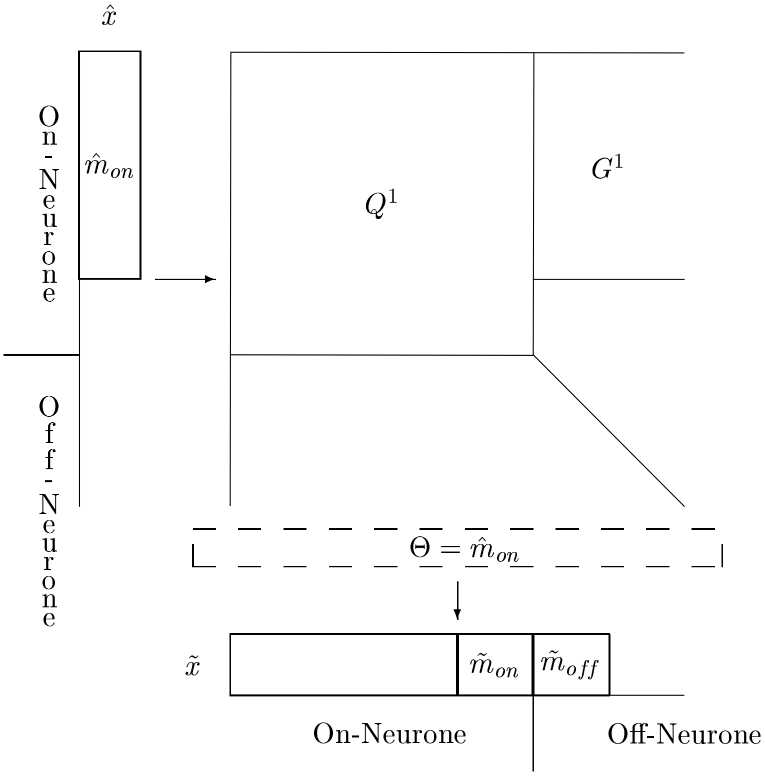

Abbildung 6.1: *Erster Retrievalschritt des Inhaltsmusters* $x = (1, 1, 1, ..., a, a, a, ...)$. Wir betrachten die linke obere Ecke der Speichermatrix. Die Adresse $\hat{x}$ hat $\hat{\alpha} = 0$ und besitzt die Aktivierung $\hat{m}_{on} = k(1 - \hat{\beta})$. Durch den Lernvorgang des Musters x wurden alle Matrixelemente im Quadrat Q^1 der Matrix auf den Wert Eins gesetzt. Da alle Spaltensummen über diesem Gebiet die Schwelle $\Theta = \hat{m}_{on}$ erreichen, sind im Ausgangsmuster alle in $\hat{x}$ fehlerhaften On-Neurone aktiviert. Die Anzahl der neu aktivierten On-Neurone ist $\tilde{m}_{on} = k\hat{\beta}$. (Die getrennte Behandlung der neuaktivierten On-Neurone gegenüber den bereits im Eingangsmuster aktiven Stellen ist für den zweiten Schritt wesentlich). Matrixelemente außerhalb des Quadrats Q^1 sind durch die Speicherung anderer Inhaltsmuster mit der Wahrscheinlichkeit $1 - p_0$ auf dem Wert Eins. Ein Off-Neuron wird überschwellig, wenn eine Spalte im Gebiet G^1 ausschließlich Einsen enthält. Die Fehlerwahrscheinlichkeit $\tilde{\alpha} = (1 - p_0)^{\hat{m}_{on}}$ gibt gerade die Wahrscheinlichkeit an, mit der im Gebiet G^1 eine Spalte mit Einsen erzeugt wurde. Die Anzahl der in diesem Schritt neu aktivierten Off-Neurone ist damit $\tilde{m}_{off} = (n - k)\tilde{\alpha}$.

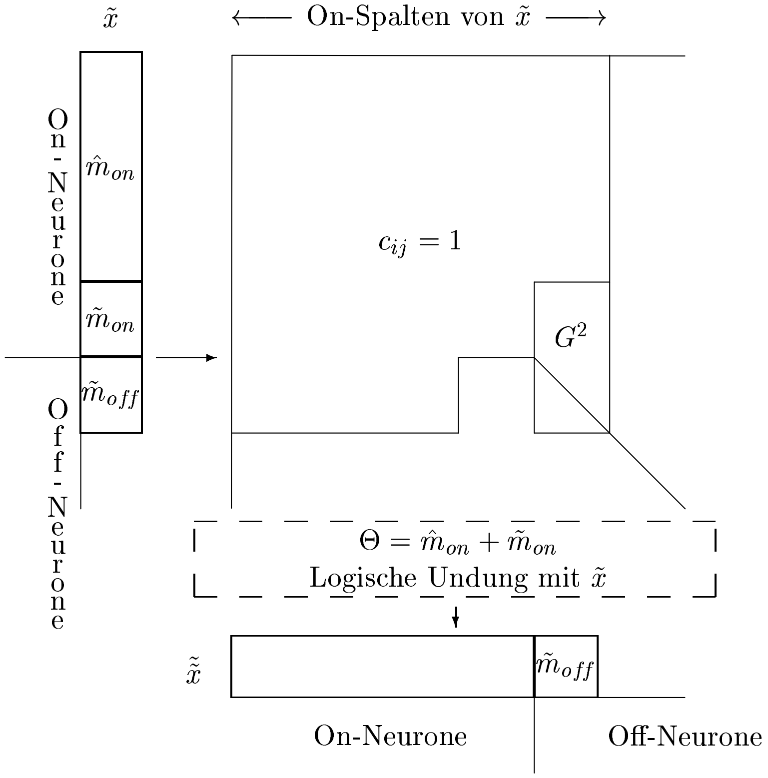

Abbildung 6.2: *Zweiter Retrievalschritt des Inhaltsmusters* $x = (1, 1, 1, ..., a, a, a, ...)$ *mit Strategie Tk+.* Als Adresse wird das Ausgangsmuster $\tilde{x}$ des ersten Retrievalschritts verwendet. Die aktivierten On-Neurone der Adresse sind eingeteilt in solche, die bereits in $\hat{x}$ aktiv waren: $\hat{m}_{on}$ und solche, die im ersten Schritt neu aktiviert wurden: $\tilde{m}_{on} = k\hat{\beta}$. Die aktiven (fehlerhaften) Off-Neurone stammen alle aus dem ersten Schritt: $\tilde{m}_{off} = (n-k)\tilde{\alpha}$. Auch können wir durch den Ausgang des ersten Retrievalschrittes bereits ein größeres Gebiet von Einskomponenten in der Matrix abstecken. Wegen der UNDungsvorschrift in Schwellenstrategie Tk+ kann lediglich ein Gebiet der Matrix das Ergebnis beeinflussen, das in den On-Spalten von $\tilde{x}$ liegt. Ein Off-Neuron wird überschwellig, falls in einer Spalte des Gebiets G^2 der Matrix mindestens $k - \hat{m}_{on}$ Einsen auftauchen. Die Wahrscheinlichkeit dafür ist $p[d^2_{off} \geq k \mid d^1_{off} \geq (1-\hat{\beta})k]$.

6.3.1 Hifi-Korrekturkapazität

Das Lernverfahren stellt den linken Abschnitt des Datenkanals in Abbildung 3.4 dar, der die Inhaltsmuster auf die Synapsenmatrix abbildet. Für diesen Kanalabschnitt wurde in [Pa92] gezeigt, daß er unabhängig vom Retrievalverfahren seine maximale asymptotische Kapazität für $p_0 = 0.5$ erreicht. Hierdurch ist bereits sichergestellt, daß asymptotisch auch bei iterativem Retrieval mit dem Optimum bei einer Beladung von $p_0 = 0.5$ gerechnet werden kann. Einschrittretrieval erzielt bei genau halb gefüllter Speichermatrix die Hifi-Korrekturkapazität von 0.17 Bit/Synapse, siehe Abschnitt 4.4.5. Dieser Wert wird bei größtmöglicher Spärlichkeit der Inhaltsmuster erreicht, mit der noch die Hifi-Bedingung erfüllbar ist; für Muster mit der Aktivierung (4.28) $p = u_1 \ln[n]/n$, wobei die Konstante u_1 die untere Schranke $v_1 = [(1-\hat\beta)\ln[2]]^{-1}$ besitzt.

Natürlich muß auch bei iterativem Retrieval die Aktivierung der Inhaltsmuster stets über dem Wert $v_0 = (\ln[2])^{-1}$ bleiben, der bei perfekter Adressierung gefordert ist (4.24). Denn falls bei perfekter Adressierung bereits im ersten Retrievalschritt das Qualitätskriterium nicht erreicht wird, kann ein Iterationsschritt (dann mit verrauschter Adressierung) nur weiter verschlechtern.

Wir machen also erneut den Ansatz $p = u_2 \ln[n]/n$ und fragen, ob bei iterativem Retrieval die Hifi-Bedingung für noch geringere Aktivität der Inhaltsmuster eingehalten werden kann, in anderen Worten, ob die Hifi-Schranke v_2 für u_2 echt unter v_1 liegt: $v_0 < v_2 < v_1$.

In Abbildung 6.2 läßt sich einfach abschätzen

$$\tilde{\tilde{\alpha}} = \tilde{\alpha}\; p[d^2_{off} \geq k \mid d^1_{off} \geq (1-\hat{\beta})k] \approx \tilde{\alpha} \binom{\tilde{m}_{on} + \tilde{m}_{off}}{\tilde{m}_{on}} (1-p_0)^{\tilde{m}_{on}}. \tag{6.2}$$

Im Bereich $\tilde{m}_{off} < \tilde{m}_{on}$ gilt die Abschätzung des Binomialfaktors $\binom{\tilde{m}_{on}+\tilde{m}_{off}}{\tilde{m}_{on}} > (2\tilde{m}_{off}/\tilde{m}_{on})^{\tilde{m}_{on}}$.

Die für die Hifi-Bedingung notwendige Bedingung (3.18) für den zweiten Schritt $\tilde{\tilde{\alpha}}/p \to 0$ kann nur erfüllt sein, wenn gilt

$$\begin{aligned} \tilde{\tilde{\alpha}}/p &= \frac{n\tilde{\tilde{\alpha}}}{u_2 \ln[n]} > \frac{\tilde{m}_{off}}{k} \left(\frac{2\tilde{m}_{off}(1-p_0)}{\tilde{m}_{on}}\right)^{\tilde{m}_{on}} \approx \exp[\tilde{m}_{on}\,(\ln[2\tilde{m}_{off}] - \ln[\tilde{m}_{on}])] \\ &\approx \exp\Big[u_2(\ln[n])^2 \Big\{1 + u_2(1-\hat{\beta})\ln[1-p_0]\Big\}\Big] \to 0 \end{aligned} \tag{6.3}$$

Daß der Exponent in (6.3) negativ wird, muß die Konstante u_2 genau die gleiche Bedingung erfüllen wie beim Einschrittretrieval, vgl. (4.28).

Durch einen zweiten Retrievalschritt ist also keine Erhöhung der asymptotischen Korrekturkapazität für Einschrittretrieval von $K = 0.17$ Bit/Synapse zu erreichen. Dieses Ergebnis läßt den Schluß zu, daß auch eine beliebige Anzahl von Retrievalschritten mit der Strategie Tk+ keine Verbesserung gegenüber Einschrittretrieval bringt, sondern genauso die asymptotische Korrekturkapazität von $K = 0.17$ Bit/Synapse erzielt wird. Die oberen Schranken aus Kapitel 5 scheinen durch Ausleseverfahren, die von einer einzigen Adresse starten – zumindest asymptotisch – nicht erreichbar zu sein. Im folgenden Abschnitt werden wir im Hinblick auf diese Fragestellung das Verhalten des endlichen Systems untersuchen.

6.3.2 Fehlerwahrscheinlichkeit im endliches System

Im endlichen System betrachten wir für eine feste Fehleranzahl j des ersten Schrittes die vollständige Binomialsumme im Gebiet G^2 von Abbildung 6.2:

$$p[d_{off}^2 \geq k \mid d_{off}^1 \geq (1-\hat{\beta})k \cap \tilde{m}_{off} = j] = \sum_{i=0}^{j} \binom{\tilde{m}_{on}-1+j}{\tilde{m}_{on}-1+i} (1-p_0)^{\tilde{m}_{on}-1+i} p_0^{j-i}. \qquad (6.4)$$

Die Zusätze von '-1' in Gleichung 6.4 kommen durch die Diagonale der Kopplungsmatrix mit $c_{ii} = 1$, die durch den mit $\tilde{m}_{off}$ adressierten Teil des Gebiets G^2 läuft und eine sichere Eins im dendritischen Potential beisteuert. Wir benutzen eine Poissonnährung für die Verteilung der Fehlerzahl im ersten Schrittes um den Mittelwert $\tilde{m}_{off}$ und erhalten damit anstelle von Gleichung (6.2):

$$\tilde{\tilde{\alpha}} = e^{-\tilde{m}_{off}} \sum_{j=1}^{n-k} \frac{\tilde{m}_{off}^{j}}{j!} \sum_{i=0}^{j} \binom{\tilde{m}_{on}-1+j}{j-i} (1-p_0)^{\tilde{m}_{on}-1+i} p_0^{j-i}. \qquad (6.5)$$

Bei der numerischen Auswertung von Gleichung (6.5) wurde noch zusätzlich eine Korrektur wegen der Abhängigkeit mehrerer Elemente einer Matrixspalte berücksichtigt. Anstatt der Terme $(1-p_0)^u p_0^v$ wurden für die Wahrscheinlichkeiten $p[\cap_{i=1}^{u} (\bar{c}_{ij} = 1) \cap_{i=u+1}^{u+v} (\bar{c}_{ij} = 0)]$ eine Abschätzung nach Proposition 4 *ii*) aus Anhang A benutzt.

Mit dieser Fehlerberechnung wurden die Theoriekurven für den zweiten Schritt in Abschnitt 6.4 erzeugt. Sie werden dort diskutiert und mit den experimentellen Werten verglichen.

6.4 Eigenschaften im endlichen System

Nachdem die asymptotische Untersuchung des zweiten Retrievalschrittes in Abschnitt 6.3 ergab, daß iteratives Retrieval gegenüber Einschrittretrieval keine verbesserten asymptotischen Kapazitätswerte erbringt, wird hier das Verhalten endlicher Speichergrößen untersucht. Speziell interessiert uns die Frage, wie schnell (für welche Systemgrößen) und von wo die asymptotischen Korrekturkapazitätswerte erreicht werden.

Anhand von Diagrammen, die einerseits aus der Analyse aus Kapitel 4 und Abschnitt 6.3, andererseits aus den Simulationsexperimenten stammen, wird das Verhalten der Modellvarianten diskutiert und abschließend verglichen.

Die Speichermodelle wurden auf einem SUN IPX Rechner simuliert. Dabei wurden die spärlichen Muster durch Pointer auf ihre 1-Komponenten dargestellt. Mit einer zufallserzeugten Menge von Inhaltsmuster, wie in Abschnitt 3.4 beschrieben, wurde der Lernvorgang ausgeführt. Dem Testschema aus Abschnitt 3.6.1 folgend, wurde aus den Adressmustern eine Menge von definiert verrauschten Testadressierungsmustern erzeugt, als Eingangsmuster im Retrieval benutzt und eine Statistik über Retrievalfehler geführt.

Bezeichnungen in diesem Abschnitt

Zum Verständnis der Bezeichnungen in den folgenden Bildern, hier noch eine tabellarische Zusammenstellung der Größen mit Erklärung:

α	Fehlerwahrscheinlichkeit in Off-Neuronen eines binren Musters
B_x	Wertebereich einer diskreten Zufallsvariable
β	Fehlerwahrscheinlichkeit in On-Neuronen eines binären Musters
f	Verhältnis falscher zu richtiger a-Komponenten in einem binären Muster
k	Anzahl von 1-Komponenten in den Inhaltsmuster
K	Korrekturkapazität
l	Anzahl der 1-Komponenten in der Adresse
M	Anzahl der gespeicherten Inhaltsmuster
μ	Aktivität eines Musters
p	Wahrscheinlichkeit, mit der eine Musterkomponente eines Inhaltsmusters On-Neuron ist
p_0	Wahrscheinlichkeit, mit der ein Matrixelement Null ist
Tk	Schwellenstrategie
$Tk+$	Schwellenstrategie
$\hat{}$	Kennzeichnung der Größen des Adressmusters
$\tilde{}$	Kennzeichnung der Größen des Ausgangsmusters
$\tilde{\tilde{}}$	Kennzeichnung der Größen des Ausgangsmusters nach dem zweiten Retrievalschritt

Tabelle 6.1: Auszug aus der Symbolliste in Anhang D

6.4.1 Binäres versus inkrementelles Lernverfahren, n=1900

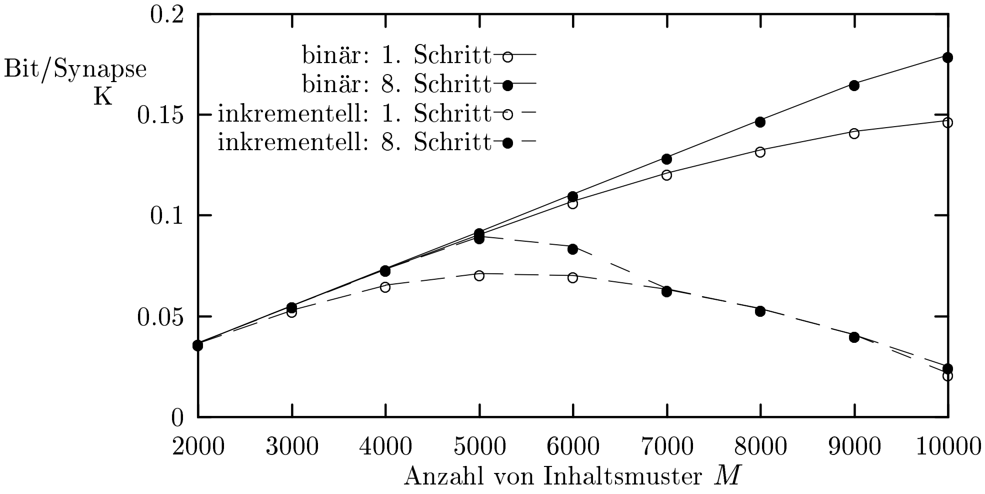

Abbildung 6.3: Korrekturkapazität für inkrementelle und binäre Speicherung, 1-ter und 8-ter Retrievalschritt. Strategie Tk+, Adressierung mit $l = 7$, $n = 1900$ und $k = 13$.

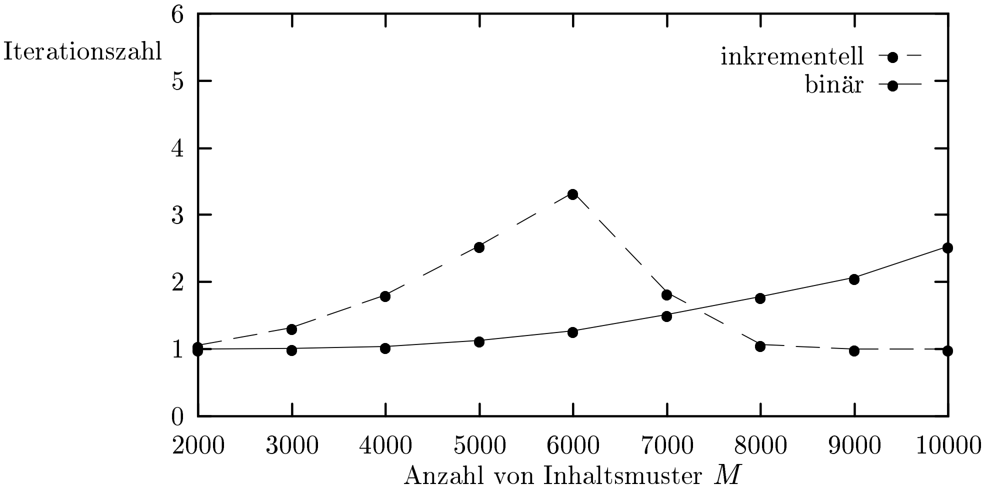

Abbildung 6.4: Iterationszeit bis zum Fixpunkt für inkrementelle und binäre Speicherung. Adressierung mit $l = 7$. $n = 1900$ und $k = 13$.

In Abbildung 6.3 ist die Korrekturkapazität für binäre, bzw. inkrementelle Speicherung über M aufgetragen. Die mit massiven Punkten gekennzeichneten Kurven erhält man für iteratives Retrieval (8-ter Schritt), die durch Kreise gekennzeichneten Kurven für Einschrittretrieval. Für den Beladungsbereich, in denen die Kapazitätskurven für iteratives Retrieval höher sind (zwischen 2000 und 7000 Mustern), steigen die Iterationszahlen bis zum Fixpunkt in Abbildung 6.4 deutlich über Eins, liegen aber immer noch sehr niedrig (< 5), vgl. hierzu die Diskussion zu Abbildung 6.7 im nächsten Abschnitt.

Wenn $M > 5000$ wird, nehmen für inkrementelle Speicherung die Retrievalfehler stärker zu, wodurch die Kapazität bald stark hinter derjenigen des binären Modelles zurückbleibt. In Abbildung 6.3 liegt das Maximum der Korrekturkapazität für binäre Speicherung gegenüber inkrementeller Speicherung bei mehr als dem doppelten Wert für Parameter l und k, die speziell für inkrementelle Speicherung optimiert wurden. Bei allen uns experimentell zugänglichen Systemgrößen (bis $n = 20000$) erweist sich binäre Speicherung als klar überlegen. Wie wir aus den asymptotischen Ergebnissen wissen – bezüglich der asymptotischen Korrekturkapazität ist die inkrementelle Speicherung leicht besser – muß sich dieses Verhalten bei sehr großen Werten von n umkehren.

Unsere Kapazitätswerte für inkrementelle Speicherung liegen auch unter den Ergebnissen in [HBFKS89, FL92], obwohl dort $p = 0.1$ also weit weniger spärliche Lernmuster verwendet wurden. Dies hat seine Ursache darin, daß dort die optimale Lernregel und $\mu = 0$, also $a < 0$ verwendet werden.

Alle Maxima der Korrekturkapazität in Abbildung 6.3 werden bei kleinen Retrievalfehlern angenommen, also in einem für Anwendungen interessanten Bereich, vgl. Abbildung 6.10.

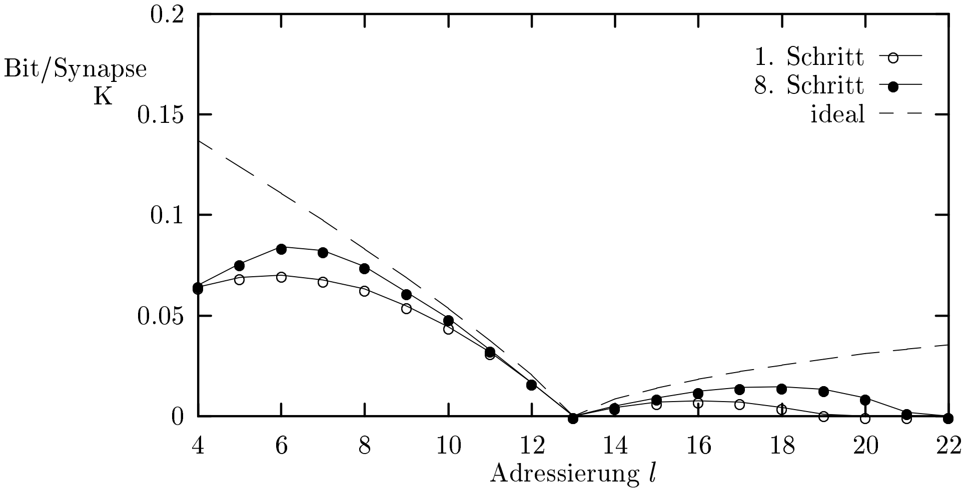

Abbildung 6.5: Korrekturkapazität für 1-ten Schritt, 8-ten Schritt und ideales Retrieval. für inkrementelle Speicherung. Adressierung für $l < 13$: $\hat{\beta} = 0$, für $l > 13$: $\hat{\alpha} = 0$. Strategie Tk+, M=6000, $n = 1900$ und $k = 13$.

Abbildung 6.5 zeigt die Korrekturkapazität für Einschrittretrieval und iterativem Retrieval für $M = 5000$, wenn l in den Eingangsmustern so variiert wird, daß immer nur ein Fehlertyp vorkommt. Die obere gestrichelte Linie ohne Messpunkte ist die fehlende Information in der verrauschten Adresse (vgl. Abbildung 3.2). Sie entspricht der Korrekturkapazität von ideal fehlerfreiem Retrieval und verschwindet für fehlerfreie Adresse, die bereits die vollständige Information über das Inhaltsmuster enthält. Die Korrekturkapazitäten stimmen in der Nähe der fehlerfreien Adresse mit der idealen Kurve überein und bleiben mit zunehmender Adressverrauschung und daraus resultierenden ansteigenden Retrievalfehlern zurück. Um so kleiner die Beladung M, umso stärker muß der Adressierungfehler sein, der zum Einsetzen von Retrievalfehlern führt. Für kleine M folgt die Messkurve also länger der (dann niedriger verlaufenden) idealen Kurve und nimmt ihre Maxima weiter außen an. Die Messkurven dieses Typs sind für alle hier untersuchten Modellvarianten M-förmig und das rechte Maximum ist gegenüber dem linken für alle Modelle kleiner. Dies liegt einerseits bereits am Verhalten der Idealkurve – sie verläuft stets nach rechts flacher – und zum anderen zeigen alle hier untersuchten Modellvarianten generell eine größere Fehlertoleranz gegenüber fehlenden Einsen in der Adresse. Im Bereich extremer Spärlichkeit, der hier untersucht wurde, wird das Optimum immer dann angenommen, wenn auschließlich On-Fehler in der Adresse vorhanden sind. Bei $p = 0.1$, also weniger spärlichen Mustern erreicht Horner et al (für inkrementelle Speicherung mit der optimalen Regel und $\mu = 0$) das Maximum für eine Adressierung mit 3/4 On-Fehler und 1/4 Off-Fehler. Es läßt sich also auch in jenem Modell die Tendenz größerer Fehlertoleranz gegenüber On-Fehlern erkennen.

6.4.2 Schwellenstrategien für binäre Speicherung, n=1900

In diesem Abschnitt vergleichen wir die in Abschnitt 6.2 eingeführten Schwellenstrategien im Modell mit binärer Speicherung.

Für die drei Schwellenstrategien ist in Abbildung 6.6 die Korrekturkapazität über M aufgetragen. Bis zum Einsetzen von Retrievalfehlern steigt die Korrekturkapazität linear mit der Musterzahl. In diesem Bereich findet fast fehlerfreies Retrieval statt. Es ließen sich mit der Adressierung $l < 6$, also größerer Adressverrauschung noch etwas höhere Kapazitätswerte erreichen. Bei Strategie Tk führt ein sehr starker Anstieg von Retrievalfehlern bereits bei kleineren M-Werten zu einem Einbruch des Kapazitätswertes. Strategie Tk+ und CA erlauben größere M-Werte und verhalten sich untereinander beinahe gleich.

In Abbildung 6.7 sind für die drei Schwellenstrategien die Iterationszeiten zum Fixpunkt über M aufgetragen. Für Werte von M, die jenseits des Kapazitätsoptimum in Abbildung 6.6 liegen, sinken die Iterationszeiten der Strategie Tk sehr schnell, später auch die von Strategie Tk+ auf einen Retrievalschritt. Diese beiden Strategien relaxieren bei Überladung gegen einen Muster, daß gegenüber den Inhaltsmuster eine höhere Aktivierung besitzt. Strategie CA verhält sich völlig anders: die Iterationszeiten nehmen für wachsendes M zu, der Zustand entfernt sich aber immer öfter vom gewünschten Muster oder endet in einem Zweierzyklus.

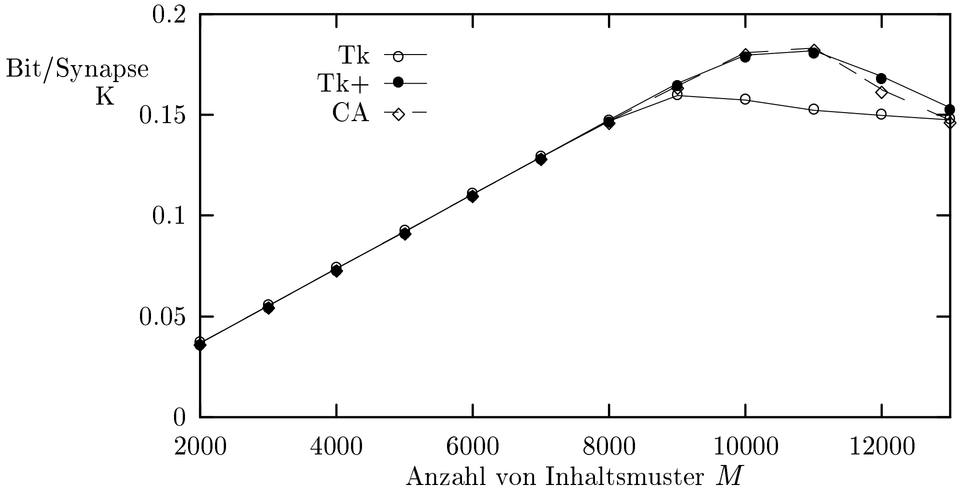

Abbildung 6.6: Korrekturkapazität der 3 Schwellstrategien nach dem 8-ten Iterationsschritt. $n = 1900$, $k = 13$ und Adressierung mit $l = 6$.

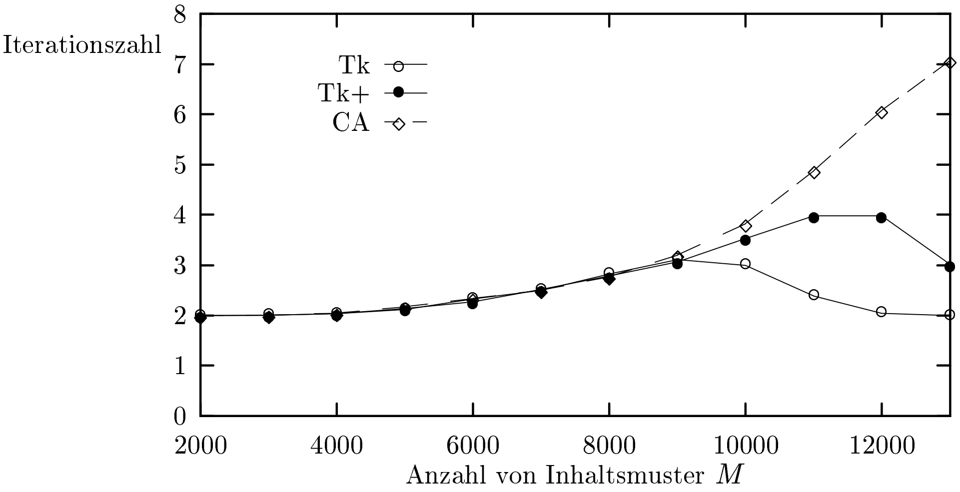

Abbildung 6.7: Iterationszeiten der 3 Strategien. $n = 1900$, $k = 13$ und Adressierung mit $l = 6$.

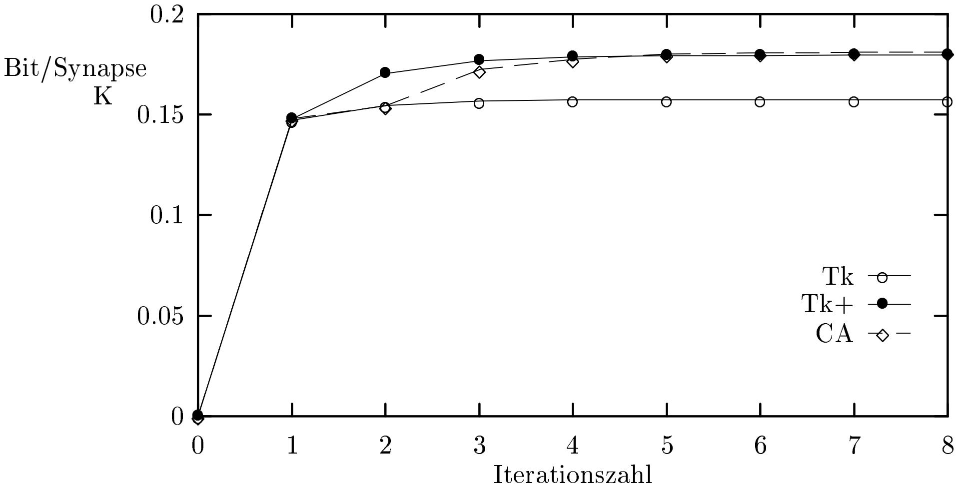

Abbildung 6.8: Korrekturkapazität für die Schwellstrategien bei binärer Speicherung und M=10000. $n = 1900$, $k = 13$, Adressierung mit $l = 6$.

Abbildung 6.8 zeigt (für fast optimale Beladung) das Zeitverhalten der Korrekturkapazität im Vergleich für die drei Schwellenstrategien. Mit Strategie Tk+ steigt die Korrekturkapazität nach dem ersten Schritt noch am stärksten weiter an, wird jedoch nach wenigen Retrievalschritten von Strategie CA eingeholt. Dagegen ist im Mittel mit Strategie Tk in den iterierten Retrievalschritten fast kein Zugewinn an Information gegenüber Einschrittretrieval zu verzeichnen. Diese Strategie ist eindeutig schlechter wegen der Explosion der Ausgangsaktivierung, die in der Iteration die bei dieser Beladung bereits häufiger auftritt. Bei der Speicheranfrage mit einem Einzelmuster führt das oft auch zu einer ganz erheblichen Verschlechterung gegenüber Einschrittretrieval. Die Iterationszeiten, für die eine Verbesserung der Korrekturkapazität zu beobachten ist, sind durchweg sehr kurz, hier $t \leq 5$ und in sämtlichen durchgeführten Experimenten immer $t \leq 10$.

In diesem Abschnitt untersuchen wir Strategie Tk+ mit binärer Speicherung, diejenige Konstellation, die sich im bisherigen Vergleich als die effektivste Modellvariante herausgestellt hat.

6.4.3 Optimale Modellvariante, n=1900

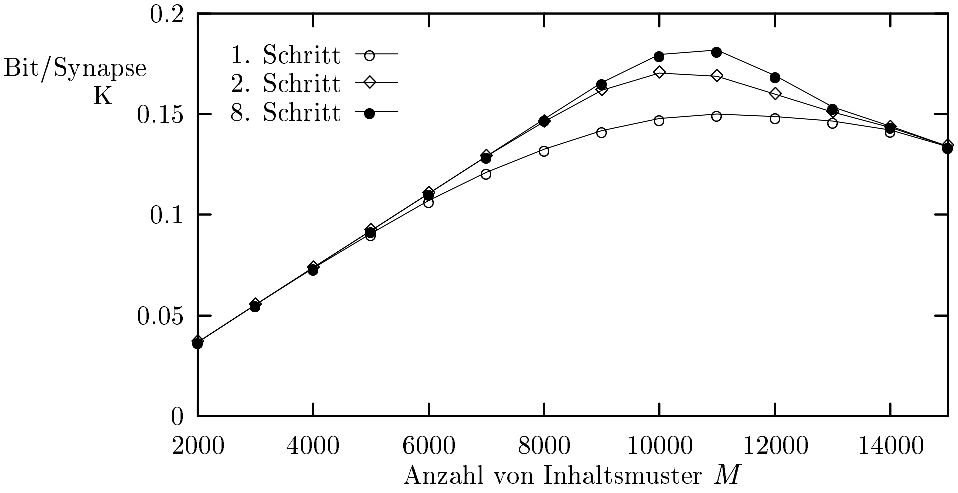

Abbildung 6.9: Korrekturkapazität für den 1-ten, 2-ten und 8-ten Schritt mit Strategie Tk+. Mit $l = 6$, $n = 1900$ und $k = 13$.

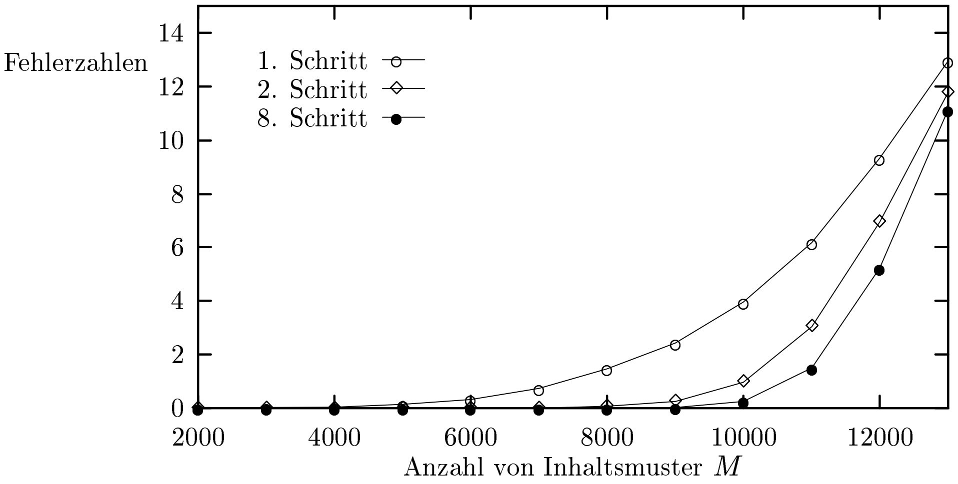

Abbildung 6.10: Fehler nach dem 1-ten, 2-ten und 8-ten Schritt. Strategie Tk+, $n = 1900$, $k = 13$ und $l = 6$.

Abbildung 6.9 gibt die Korrekturkapazität im ersten und zweiten Retrievalschritt, sowie nach Erreichen des Fixpunkts (8-ten Schritt) wieder. Hier ist, wie schon in Abbildung 6.3, das Fenster im Beladungsbereich deutlich zu sehen, in welchem iteratives Retrieval eine Erhöhung der Kapazität bewirkt. Iteratives Retrieval verbessert das Ergebnis erst bei hohen Beladungen, wo Einschrittretrieval bereits mit mittleren Fehlerzahlen über Eins (vgl. Abbildung 6.10) arbeitet. Ziemlich genau am Optimum für Einschrittretrieval erreicht iteratives Retrieval eine weitere Erhöhung der Korrekturkapazität um ca. 20 %. Für extrem hohe Beladungen geht die Kapazitätskurve wieder auf die Einschrittretrievalkurve zurück, die Fehler lassen sich dann durch weitere Retrievalschritte nicht mehr reduzieren.

In Abbildung 6.10 sind die mittleren Fehlerzahlen aufgetragen. Bei der Beladung $M = 11000$, bei der die Korrekturkapazität von iterativem Retrieval in Abbildung 6.9 maximal wird, erreicht Einschrittretrieval das Maximum bei einem mittleren Fehlerzahl von beinahe $k/2$. Hingegen wird das Kapazitätsoptimum von iterativem Retrieval bei wesentlich geringeren Fehlerzahlen von ungefähr $k/12$ angenommen. Für Anwendungen, die eine Fehlerschranke vorgeben, kann man also bei iterativem Retrieval mit der Beladung näher an den optimalen Wert gehen als bei Einschrittretrieval und damit einen größeren Zugewinn an Korrekturkapazität erzielen als es die Differenz der Optima in Abbildung 6.9 erwarten läßt.

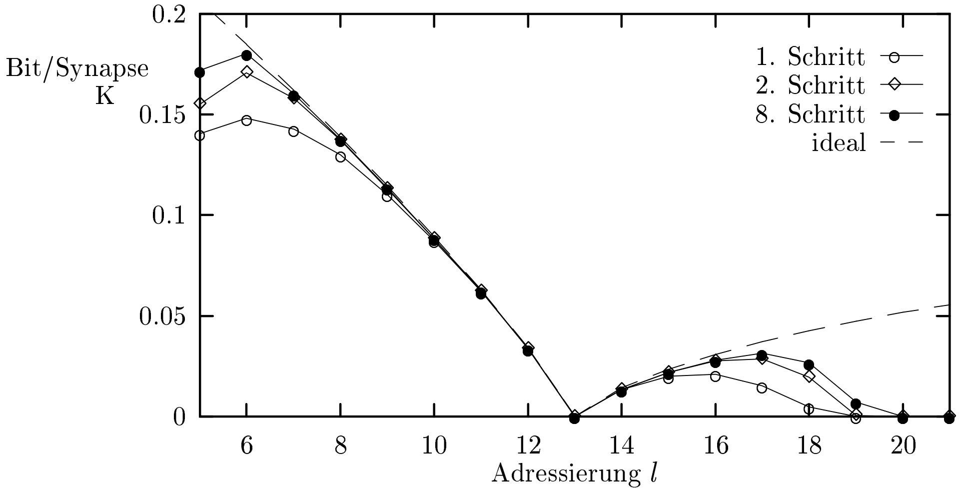

Abbildung 6.11: Korrekturkapazität für 1-ten und 8-ten Retrievalschritt mit Strategie Tk+ und ideales Retrieval bei M=10000. $n = 1900$ und $k = 13$.

Die Darstellung in Abbildung 6.11 ist analog zu Abbildung 6.5 und gibt wieder, in welcher Weise die Korrekturkapazität von der Adressierung abhängt. Maximale Kapazität für $k = 13$ erhält man durch Adressierung mit $l = 6$. Abbildung 6.12 zeigt die Korrekturkapazität für variiertes k. Im aufgetragenen Bereich von k erhält man beinahe konstante Ergebnisse, die jedoch für ungerades k leicht höher liegen als für geradezahliges. Diese Schwankungen sind eine Folge davon, daß bei endlicher Systemgröße der ganzzahlige Parameter l nicht beliebig fein adjustiert werden kann. Bei dieser Systemgröße stellt sich eine Adressierung mit l leicht unter $k/2$ als optimal heraus, die für ungeradzahliges k mit $l = k/2 - 1$ besser erreicht wird. Wie wir aus der Theorie wissen, vgl. (4.30), ist asymptotisch die Adressierung $l = k/2$

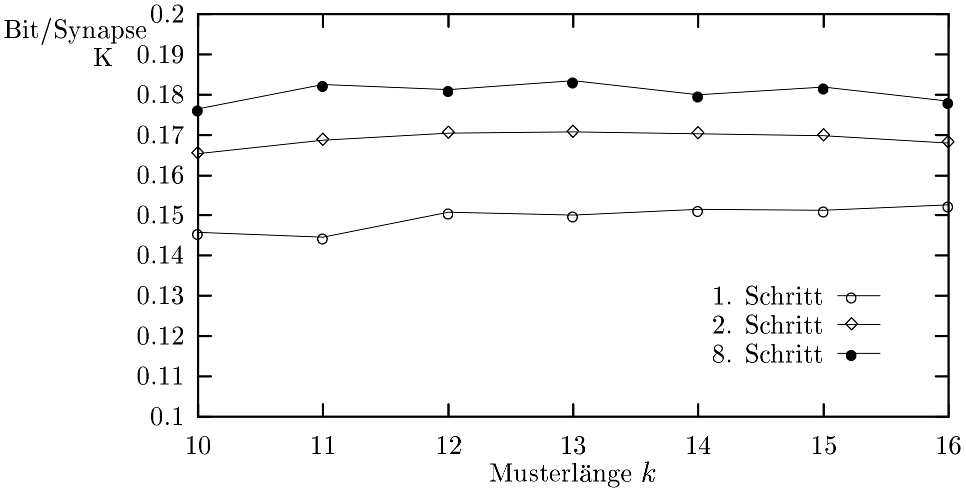

Abbildung 6.12: Korrekturkapazität für 1-ten 2-ten und 8-ten Schritt in Abhängigkeit von k für $n = 1900$. Die Addresse enthält $l =$ Integer$(k/2)$ Einsen, die Beladung M ist für maximale Kapazität des 8-ten Schrittes optimiert.

optimal.

6.4.4 Theorie und Experiment am Optimum

Dieser Abschnitt enthält für das optimale Modell einen Vergleich zwischen Theorie und den experimentellen Ergebnissen im Arbeitsbereich mit maximaler Korrekturkapazität, d.h. mit dafür optimierten Parametern M, k und l. Die Diagramme stellen die verschiedenen Größen in Abhängigkeit von der Matrixgröße, also von n^2 dar.

In Abbildung 6.13 ist die Anzahl der Adresseinsen l am Optimum aus der endlichen Theorie mit Simulationspunkten aufgetragen. Der von der asymptotischen Theorie vorrausgesagte Verlauf: $l = \log_2[n]$ (nach (4.23),(4.28)) ist gestrichelt eingetragen, er liegt höher. Die optimalen Inhaltsmuster im endlichen System besitzen demnach eine niedrigere Aktivierung $k \approx 2l$ als man asymptotisch erwarten würde. Für große n gleichen sich die Steigungen beider Kurven an und der Offset wird klein gegenüber l und somit vernachlässigbar für $n \to \infty$.

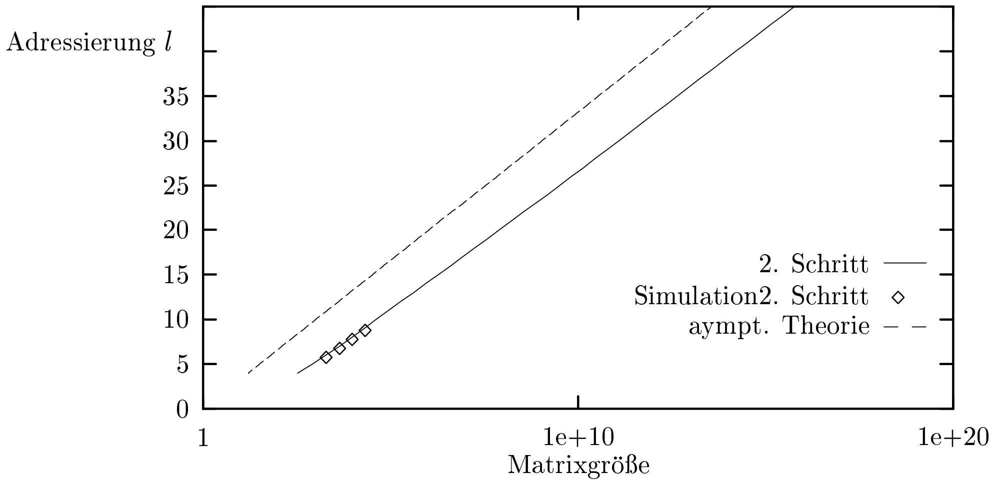

Abbildung 6.13: Optimales l für 2-ten Retrievalschritt vs. Matrixgröße (logarithmisch). Theorie und Experiment.

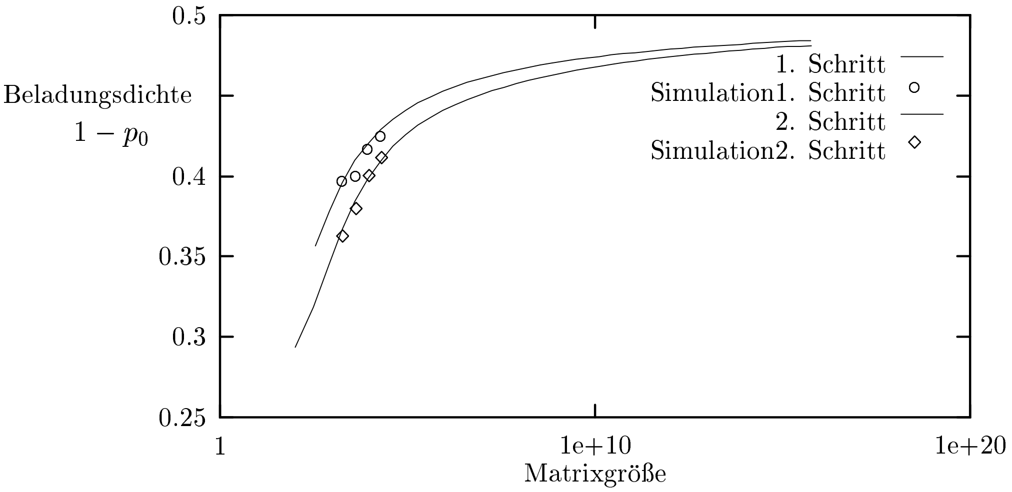

Abbildung 6.14: Optimale Beladungsdichte p_0 vs. Matrixgröße (logarithmisch).Theorie und Experiment.

Wie in Abbildung 6.14 zu sehen ist, führt die kleinere Aktivierung der Inhaltsmuster in Abbildung 6.13 dazu (und wird auch nicht durch eine höhere Musterzahl M kompensiert), daß der beste Beladungswert der Matrix $1 - p_0$ im endlichen System noch deutlich unter dem asymptotisch optimalen Wert von 0.5 liegt. Der asymptotische Wert wird erst weit außerhalb des in Abbildung 6.14 aufgetragenen Bereichs der Systemgröße gut approximiert.

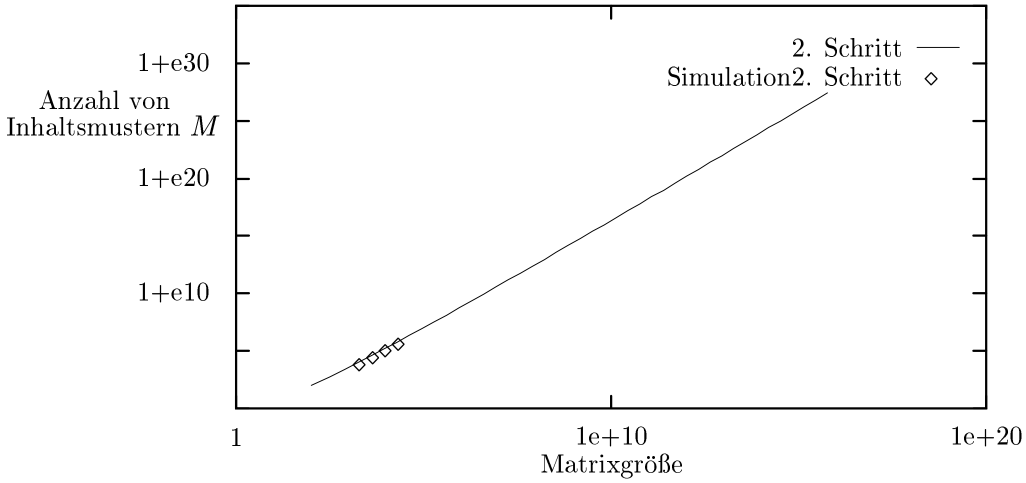

Abbildung 6.15: Optimale Musterzahl (logarithmisch) vs. Matrixgröße (logarithmisch). Theorie und Experiment.

Abbildung 6.15 zeigt den Verlauf der optimalen Musterzahl gegen die Matrixgröße, einen beinahe linearen Verlauf, also $M \propto n^2$. Für spärliche Muster sind also schon für 'simulierbare' Systemgrößen weit mehr Muster speicherbar, als das Netzwerk Neurone enthält.

Das Fehlerverhältnis für den ersten und zweiten Schritt in Abbildung 6.16 geht asymptotisch gegen Null. Für den zweiten Schritt wird der Grenzwert sehr viel schneller erreicht, für höhere Iterationszahlen (hier nicht eingetragen, vgl. Abbildung 6.10) noch schneller. Die Fehlerzahl ist im Vergleich zu den Größen aus den Abbildungen 6.13 bis 6.15, wo Experiment und Theorie gut übereinstimmen, eine sehr sensible Größe. Besonders für den zweiten Schritt und kleine Systemgröße ist die Übereinstimmung mit der Theorie schlecht, die theoretischen Werte sind etwas zu optimistisch. Wir vermuten, daß diese Abweichung eine Folge davon ist, daß in der Fehlerberechnung in Abschnitt 6.3 die statistische Abhängigkeit unterschiedlicher Spalten vernachlässigt wird, was für kleine Systemgröße unzulässig ist. Eine einfache Modifikation unserer Methode zur Berücksichtigung dieser Effekte ist jedoch nicht möglich.

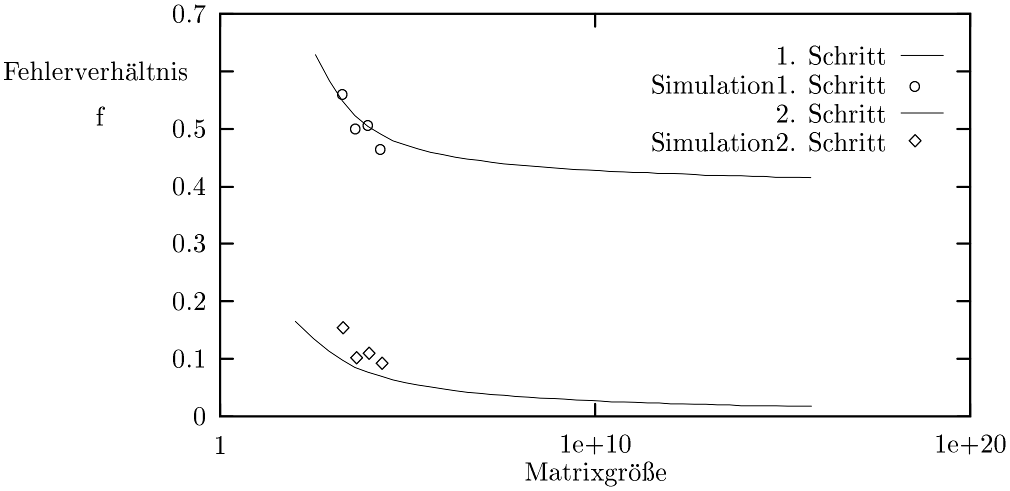

Abbildung 6.16: Fehlerverhältnis am Optimum für den ersten und zweiten Schritt vs. Matrixgröße (logarithmisch). Theorie und Experiment.

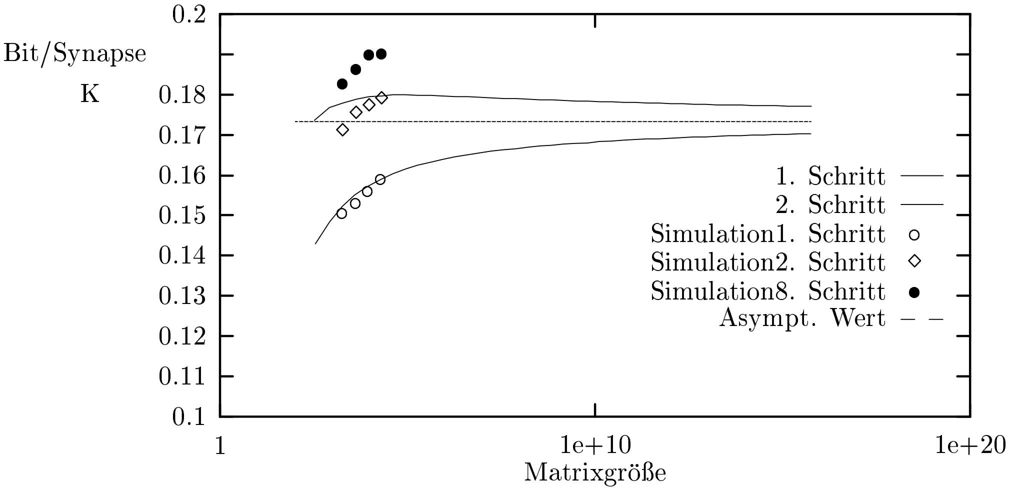

Abbildung 6.17: Optimale Korrekturkapazität für den 1-ten, 2-ten Retrievalschritt und iteratives Retrieval bis zum Fixpunkt vs. Matrixgröße (logarithmisch), Theorie und Experiment für optimiertes k und l und M. Asymtotischer Wert: gestrichelte Waagerechte.

Abbildung 6.17 zeigt das Verhalten der optimalen Korrekturkapazität für wachsendes n^2. Während sich die Kapazität für Einschrittretrieval dem asymtotischen Wert von unten nähert, erreicht der zweite Schritt in Theorie und Experiment bei endlichen Systemgrößen Werte über dem asymptotischen Wert. Für iteratives Retrieval bis zum Fixpunkt[3] sind nur die experimentellen Werte eingetragen, die Werte sind noch deutlich höher als für den 2-ten Schritt.

Wie bereits in Abbildung 6.16 erkennt man auch in diesem Diagramm, daß die Theorie bei kleinem n für den 2-ten Retrievalschritt zu optimistisch ist (vgl. hierzu die Diskussion zu Abbildung 6.16). Für größeres n wird jedoch die Übereinstimmung zwischen Theorie und Experiment besser.

Der Verlauf der Theoriekurve für den zweiten Schritt und das asymptotische Ergebnis, daß für alle Retrievalschritte der gleiche asymptotische Wert gilt, läßt auf ein bemerkenswertes Verhalten der Kapazitätskurven für iteratives Retrieval schließen: Die optimalen Kapazitätswerte werden für endliche Systemgröße oberhalb des asymptotischen Wertes angenommen. Für weiter wachsendes n findet ein leichtes Absinken der Kapazität statt und schließlich wird sehr weit draußen der asymptotische Wert von oben angenähert.

Der letzte experimentelle Punkt für Fixpunktretrieval bei $n = 20000$ scheint ungefähr das Maximum der Kurve zu repräsentieren. Mit der zur Verfügung stehenden Rechenleistung war es nicht möglich, mit Experimenten in den fallenden Bereich der Kurve vorzustoßen. Der extrem flache Abfall ist experimentell auch schwer von einer ungenügenden Optimierung der Parameter zu diskriminieren, die für große Systeme sehr rechenzeitaufwendig und nur in weniger feinen Schritten durchführbar ist.

Die Abbildungen 6.16 und 6.17 zeigen, daß iteratives Retrieval gegenüber Einschrittretrieval bei endlicher Systemgröße ein ca. 20% höheres Optimum der Kapazität aufweist. Gerade für Anwendungen, die nur geringe Retrievalfehler erlauben, kann wegen der niedrigeren Retrievalfehler von iterativem Retrieval (1/5 gegenüber Einschrittretrieval) bei gleichem Qualitätskriterium wesentlich höher beladen und somit durch iteratives Retrieval eine Steigerung der Kapazität erreicht werden, die weit über 20% liegt.

---

[3] Eigentlich sind es Ergebnisse für Retrieval bis zum 8-ten Schritt, wegen der kurzen Iterationszeiten gleichbedeutend mit Fixpunktretrieval.

Kapitel 7

Diskussion und Zusammenfassung

Für die Analyse von rückgekoppelten neuronalen Netzmodellen haben sich in der Literatur unterschiedliche Methoden etabliert. Nur ein sehr kleiner Teil der Arbeiten, so [G-M76, Ai88, Ai89, GR91], benutzt elementare statistische Methoden ähnlich unserer Vorgehensweise in Kapitel 6.

In einer der ersten Arbeiten über binäre Speicherung von endlichen spärlichen Mustern von Gardner-Medwin [G-M76] werden die iterierten Retrievalschritte gleich behandelt wie der erste Schritt und die Abhängigkeit zwischen Synapsen- und Eingangs-Noise schlichtweg vernachlässigt. Die eng daran anknüpfende Arbeit von Gibson und Robinson [GR91] berücksichtigt diese Korrelationen, es wird allerdings eine andere Schwellenstrategie benutzt und die Binomialverteilung in den Gleichungen (4.11) und (4.12) durch die Gaussverteilung approximiert, eine Näherung, die in dem hier untersuchten Bereich extremer Spärlichkeit ungenau wird.

Der weitaus größte Teil von Arbeiten verwendet Methoden, die aus der Thermodynamik stammen. Im nächsten Abschnitt werden die bekanntesten dieser Verfahren zur Analyse des Retrievalvorgangs kurz besprochen.

7.1 Thermodynamische Methode

7.1.1 Energiefunktion

Bei Autoassoziation erzeugen lokale Lernregeln im Lernvorgang symmetrische Synapsenmatrizen, d.h.: $c_{ij} = c_{ji}$. Es kann dann immer eine Energiefunktion[1] als Zustandsgröße für das rückgekoppelte System angegeben werden, die durch jeden Updateschritt gemäß der Vorschrift (2.2) monoton erniedrigt wird. Die Zustände an den Minima der Energiefunktion sind die *Fixpunkte* des Systems, denn wegen der Monotonieeigenschaft kommen die durch die Updatevorschrift hervorgerufenen Zustandsänderungen zum Stillstand, sobald das System einen dieser Fixpunkte erreicht hat. Durch die Energiefunktion, die natürlich mitbestimmt ist von der Schwellenstrategie, ist das Retrievalverhalten festgelegt. Zwei Modelle mit derselben Energiefunktion (modulo Faktoren) verhalten sich daher äquivalent.

Bei *Fixpunktretrieval* – wenn die Rückkopplung im Retrievalprozess immer bis zu einem solchen Grenzzyklus aufrechterhalten wird – sind die Ausgangsmuster also vollständig durch die Minima der Energiefunktion festgelegt.

---

[1] oft auch Liapunov-Funktion oder Kostenfunktion genannt.

Bei der Schwellenstrategie CA ist zu beachten, daß im Gegensatz zu den beiden anderen hier betrachteten Strategien außerdem die Möglichkeit besteht, daß der Retrievalprozess nicht in einem stationären Fixpunkt mündet, sondern bei hin- und herspringendem Schwellenwert in eine Folge von zwei zyklisch sich abwechselnde Systemzustände, einen Zweierzyklus[2] gerät. Solche Ausgangszustände sind bei Mustervervollständigung nicht direkt interpretierbar, in einigen Modellen werden Zweierzyklen als Zurückweisung des verwendeten Eingangsmusters interpretiert.

Konstruktion der Energiefunktion

Sei zunächst $c_{ii} = 0 \; \forall i$. Für die Zustandsänderung einer Komponente im Aktivitätsmuster $z_i \to \bar{z}_i$ soll das Vorzeichen der dadurch hervorgerufenen Energieänderung im System $\triangle \mathcal{H}_i$ immer dann negativ werden, wenn die Update-Gleichung (2.2) die Änderung vorschreibt. Der Betrag von $\triangle \mathcal{H}_i$ soll proportional zum Abstand vom dendritischen Potential zur Schwelle sein. Wir setzen also für die Energiedifferenz zweier Zustände $z^1 := (z_1, ..., z_i, ..., z_n)$ und $z^2 := (z_1, ..., \bar{z}_i, ..., z_n)$

$$\begin{aligned} \triangle \mathcal{H} &= \mathcal{H}(z^2) - \mathcal{H}(z^1) := -(d_i - \Theta)\mathrm{sgn}(z_i) \\ &= -\frac{1}{2}(d_i - \Theta)\left\{\mathrm{sgn}(\bar{z}_i) - \mathrm{sgn}(z_i)\right\} \end{aligned} \tag{7.1}$$

hier mit $\mathrm{sgn}(0) := -1$.

Die folgende Definition der Energiefunktion als Zustandsgröße des Systems ist mit Ansatz (7.1) kompatibel[3]:

$$\begin{aligned} \mathcal{H}(z) &:= -\frac{1}{2}\sum_j (d_j - \Theta)\mathrm{sgn}(z_j) \\ &= -\frac{1}{1-a}\left\{\sum_j (d_j - \Theta) z_j + \frac{1+a}{2}(\sum_j d_j - n\Theta)\right\} \end{aligned} \tag{7.2}$$

Sie bewertet für ein Sollmuster z mit gleicher Gewichtung an a und 1 Komponenten die Übereinstimmung zwischen Neuronenzustand und dendritischem Potential.

Mit einer Schwellenwahl $\Theta = E(\frac{1}{n}\sum_j d_j)$ wird der erste Term in (7.2) dominant und es verbleibt

$$\mathcal{H}(z) = -\frac{1}{1-a}\left\{\sum_j d_j z_j - \mu(z)\sum_j d_j\right\}, \tag{7.3}$$

mit der Definition der Aktivität $\mu(z)$ aus (2.4). Diese Energiefunktion besitzt zwar für Modelle mit beliebigen a-Werten dieselbe Form, für allgemeine Wechselwirkungsmatrizen verhalten sich Modelle mit verschiedenen a-Werten jedoch nicht gleich.

Spinregel

Mit der Spinregel-Matrix führt obige Schwellenwahl auf $\Theta = 0$ und damit zum Verschwinden des zweiten Terms in Gleichung 7.3.

Im Spinmodell, d.h. für $a = -1$ erhalten wir die Energiefunktion des Spinglasmodells

$$\mathcal{H}(z) = -\frac{1}{2}\sum_{ij} c_{ij} z_i z_j. \tag{7.4}$$

---

[2] wie man es beim Modell mit parallelem Update kennt, vgl. Peretto [Pe88].

[3] mit $\mathrm{sgn}(x) = [2x - (1-a)]/(1-a)$ für $x \in \{a, 1\}$.

Für $a = 0$ entspricht (7.4), lediglich mit anderem Vorfaktor, der von Hopfield [Hd82] vorgeschlagenen Energiefunktion. Für diese Lernregel verhalten sich also Spinmodell und (a=0)-Modell äquivalent.

Hebbregel
Bei der Matrix, die mit der Hebbregel erzeugt wurde, ist wegen $c_{ij} \geq 0$ immer $\Theta > 0$ optimal. Man erkennt schon in (7.4), daß dann der Energiewert eines Zustands von $\mu(z)$ abhängig ist und daher keine Isomorphie zwischen den Modellen mit unterschiedlichen a-Werten bestehen kann. Bei $a = 0$ erhält man

$$\mathcal{H}(z) = -\sum_{\nu} \left[< x^{\nu}, z >^2 + n\mu(z)\mu(x^{\nu}) < x^{\nu}, z > \right] =: -\sum_{\nu} E_{\nu}, \tag{7.5}$$

$< x, y > := \sum_{i=1}^{n} x_i y_i$ ist das Skalarprodukt von x und y. Das Skalarprodukt zwischen einem Lernmuster x^{ν} und dem Systemzustand z wird im Isingmodell oft *Überlapp* genannt: $m^{\nu} := < x^{\nu}, z >$. Für $a = -1$ ist der Überlapp äquivalent zur Hammingdistanz $m^{\nu} = 1 - \frac{2}{n} h(x^{\nu}, z)$, der Überlapp gibt die Anzahl übereinstimmender Komponenten an. Anders für $a = 0$, hier gibt m^{ν} die Anzahl übereinstimmender On-Komponenten der Muster an und legt die Hammingdistanz nicht eindeutig fest: $m^{\nu} = \frac{1}{2} \left[\mu(x^{\nu}) + \mu(z) - \frac{1}{n} h(x^{\nu}, z) \right]$.

Für allgemeine Lernregeln, die von der Spinregel abweichen, verhalten sich Modelle mit unterschiedlichem a-Wert nicht äquivalent.

Binäre Hebbregel
Für binäre Speicherung wird die Energiefunktion schlecht handhabbar, da sie wegen der nichtlinearen Operation im Lernprozess nicht mehr in eine Summe von M unabhängigen Beiträgen E_{ν} für jedes Lernmuster zerfällt. Um der Vollständigkeit willen schreiben wir die Energiefunktion an dieser Stelle für $a = 0$ einmal hin, ohne sie jedoch weiter zu verwenden

$$\mathcal{H}(z) = -\sum_{\nu} \left\{ E_{\nu} - \sum_{\mu \neq \nu} \left\{ E_{\nu\mu} - \sum_{\mu \neq \nu, \eta \neq \nu, \mu} \{ E_{\nu\mu\eta} - \ldots \} \right\} \right\} + (-1)^M E_{12\ldots M} \tag{7.6}$$

mit der Definition

$$E_{\nu\mu} := < x^{\nu} x^{\mu}, z >^2 + n\mu(z)\mu(x^{\nu} x^{\mu}) < x^{\nu} x^{\mu}, z > .$$

Die Produktschreibweise $x^{\nu} x^{\mu}$ bezeichnet das Muster, das durch logische Undung der beiden Muster x^{μ} und x^{ν}, im $(a = 0)$-Modell durch komponentenweise Multiplikation entsteht. Ganz analog werden in den Termen $E_{\nu\mu\eta}$ bis $E_{12\ldots M}$ höhere Produkte mit drei bis zu M Mustern verwendet.

7.1.2 Spinglas-Theorie

Die theoretische Beschreibung des rückgekoppelten neuronalen Netzwerks im sehr hochdimensionalen Zustandsraum der Neurone ist weder durchführbar noch relevant. Im Spinmodell (d.h. $a = -1$ und Spinregel) wird stattdessen immer das Verhalten der Überlapps zwischen Systemzustand mit den Inhaltsmustern als die *Ordnungparameter* des Systems beschrieben. Ordnungsparameter wurden in thermodynamischen mikroskopisch atomaren Vielteilchenmodellen für Gase und Festkörper[4] in der Physik benutzt, um makroskopische Eigenschaften wie Druck oder Magnetisierung zu beschreiben.

---

[4] Das dendritische Potential eines Neurons entspricht dem Feld an einem Gitteratom, der Neuronenzustand dem atomaren Spin.

Das kollektive neuronale Verhalten wird auf ein repräsentatives Neuron in einem mittleren dendritischen Potential reduziert, auf ein sogenanntes Mean-Field-Problem. Das mittlere Feld hängt über die Rückkopplung wieder selbst vom Zustand des repräsentativen Neurons ab. Die Werte der Ordnungsparameter in den Gleichgewichtszuständen des Systems lassen sich – für konstante Dichten im Limes divergierender Systemgröße, dem *thermodynamischer Limes* (der unseren asymptotischen Rechnungen entspricht) – durch Selbstkonsistenzlösung des Mean-Field-Problems ermitteln.

Die Mean-Field Theorie, mit der Fixpunktretrieval im Assoziativspeicher von Amit et al [AGS87] untersucht wurde, stammt ursprünglich von der Behandlung der Spinglässer[5].

Die Spinglas-Methode liefert asymptotische Aussagen über die Fehlerwahrscheinlichkeiten der stabilen Zustände, der Fixpunkte des Systems, woraus sich, in hier definierten Termini, für vorgegebenes Qualitätskriterium die Musterkapazität P und die Ausgangskapazität A ergibt. Übersichtsdarstellungen dieser Methode finden sich mittlerweile in zahlreichen Büchern und Artikeln, so in [At89, HKP91, DHS91, Bu88, Gu90, Kü91].

Diese Methode wurde auch auf Modelle mit Parallel-update [FK88] und auf spärliche Muster [TF88, P-VA89, BDS89] und binäre Speicherung [GRS90] erweitert. Für spärliche Muster muß noch zusätzlich der Limes $p \to 0$, der spärliche Limes vollzogen werden.

7.1.3 Analyse der Dynamik

Die Spinglas-Theorie als eine Methode der Gleichgewichtsthermodynamik läßt keine Aussagen zu über das transiente Verhalten des Systems, ausgehend von einem Adressmuster. Die Beurteilung der Mustervervollständigungseigenschaft erfordert eigentlich die Auslotung der Attraktionsbasins der Inhaltsmuster, also solcher Gebiete im Zustandsraum, von denen aus ein Inhaltsmuster wieder vervollständigt werden kann. Zur Bestimmung der Korrekturkapazität ist eine Betrachtung des dynamischen Geschehens notwendig, wie sich während des iterierten Retrievalprozesses, ausgehend von der verrauschten Adresse, der Systemzustand nach und nach ändert.

Eine Analyse der ersten zwei Retrievalschritte wurde mit in [GDM87, HO90/1, HO90/2] mit der Methode erzeugender Funktionale gerechnet. Für größere Zeiten wird in [HO90/1, HO90/2] für eine nichtlokale Lernregel[6] eine approximative Theorie entwickelt. Inkrementelle Speicherung mit der optimalen Lernregel für $\mu = 0$ wird von Horner et al in [HBFKS89] mit einer dynamischen Mean-Field Theorie untersucht.

7.1.4 Die Gardner Methode

E. Gardner berechnete in [Ga88] eine obere Schranke P_G der Musterkapazität, die durch beliebige Lernregeln erzielbar ist[7]. Es wird im hochdimensionalen Raum der durch die Bedingung $\sum_j c_{ij}^2 = N \; \forall \; i$ normierten synaptischen Gewichte das Gebiet untersucht, in dem Zwangsbedingungen für die Lernmuster erfüllt sind, sogenannte *Einbettungsbedingungen*, welche die

---

[5] Spingläser sind magnetische Festkörper mit konkurrierenden para- und antiparamagnetischen Kopplungen zwischen den quantisierten magnetischen Momenten der Gitteratome. Hierdurch sind eine sehr große Anzahl von ungeordneten (Glas-artig ungeordneten) Spinkonfigurationen stabil. Die ungeordneten magnetischen Wechselwirkungen wurden als Gaussverteilte Zufallsvariablen modelliert. Durch einen geschickten Rechentrick, die Replikamethode, ist es möglich [KS78], dafür eine Mean-Field Theorie mit Selbstkonsistenzgleichungen zu entwickeln.

[6] die Pseudoinverse-Lernregel

[7] Mit geometrischen Argumenten wurde die Schranke P_G bereits wesentlich früher von Cover [Co65] bestimmt.

Stabilität jedes Musters fordern:

$$x_i^\nu \sum_j c_{ij} x_j^\nu \geq \kappa\sqrt{N} \ \ \forall i, \nu. \tag{7.7}$$

Das Volumen dieses Gebietes erhält man durch eine Integration, gemittelt über die möglichen Sätze von Lernmustern, welche ebenfalls mit der Replikamethode ausgeführt werden kann. Die Musterkapazität P_G bei der Beladung, für die dieses Integral gerade verschwindet, stellt das Maximum dar, für das prinzipiell noch synaptische Kopplungen gefunden werden können.

Mit den Zwangsbedingungen (7.7) kann wieder eine Energiefunktion, jetzt im Raum der synaptischen Kopplungen formuliert werden. Gardner zeigt weiter, daß für Beladungen unterhalb P_G durch die Perceptron Lernregel, vgl. [Ro62], in iterativen Lernprozeduren die synaptischen Kopplungen c_{ij} an den Minima der Energiefunktion nach endlichen Lernschritten gefunden werden. Der Vergleich der Gardner Schranke mit der Musterkapazität für lokale Lernregeln erlaubt eine Aussage darüber, ob Perceptron-artige Lernregeln, die zusätzlich ein Sollsignal auswerten (lokale Dreitermlernregeln), oder nichtlokale Lernregeln eine zusätzliche Verbesserung gegenüber lokalen Zweitermlernregeln erbringen. Diese Methode wurde in [GS90] auf binäre Synapsen erweitert.

7.2 Autoassoziation im Assoziativspeicher

Musterkapazität
Dieser Abschnitt gibt einen Überblick des Wissens über Autoassoziation, wie es einerseits aus den besprochenen Literaturquellen bekannt ist und andererseits in dieser Arbeit gewonnen wurde.

7.2.1 Asymptotische Ergebnisse

In [AGS87] wurde für nichtspärliche Muster mit $p = 0.5$ im Hopfieldmodell eine Musterkapazität von $P = 0.14$ Bit/Synapse bestimmt. Bei extensiver Beladung, also $P > 0$, behalten die Fixpunkte nichtverschwindende Abweichungen von den Inhaltsmustern, die Hifi-Bedingung ist für nichtspärliche Muster generell nicht erreichbar[8]. Der Wert der Gardner Schranke in [Ga88] von $P_G = 2$ Bit/Synapse zeigt, daß bei Speicherung nichtspärlicher Muster komplexe Lernregeln gegenüber lokalen Zweitermlernregeln große Verbesserungen bewirken.

Für spärliche Lernmuster wurde mit einer inkrementellen Lernregel, die gegen die Hebbregel konvergiert, in [TF88] die sehr viel höhere Musterkapazität von $P = 0.72$ Bit/Synapse bestimmt. Auch für Einschrittretrieval in Abschnitt 4.5 und durch das Fixpunkt-Ausleseverfahren aus Kapitel 5 wird dieser Wert reproduziert. Das bedeutet, daß die asymptotische Musterkapazität von Einschrittretrieval mit der inkrementellen Hebbregel nicht durch aufwendigere Ausleseverfahren verbessert werden kann. Ferner fällt dieses Ergebnis mit der Schranke von E. Gardner [Ga88] zusammen. Für spärliche Inhaltsmuster kann auch die Verwendung nichtlokaler Lernregeln keine weitere Verbesserung der asymptotischen Musterkapazität hervorrufen.

Für binäre Speicherung finden wir in der Literatur eine Inkonsistenz: die Gardner Schranke $P_G = 0.29$ Bit/Synapse aus [GS90] fällt weit niedriger aus als die Musterkapazität für binäre Speicherung mit $P = 0.69$ Bit/Synapse. Dies wurde bereits in [Na91] aufgegriffen und mit der konstanten Aktivierung der Inhaltsmuster in der Willshaw Rechnung erklärt. Mit der

---

[8] die gleiche Aussage gilt auch für das Einschrittretrievalmodell, wie in Kapitel 4 gezeigt wurde.

Fixpunkt-Ausleseverfahren konnten wir hier auch Inhaltsmuster mit nur im Mittel konstanter Aktivierung behandeln, wobei dann ebenfalls das Willshaw Resultat reproduziert wird. Dieser Sachverhalt widerspricht der Erklärung aus [Na91]: die Inkonsistenz scheint also doch eher durch die unterschiedlichen Qualitätskriterien erklärbar. Während die Einbettungsbedingung (7.7) in der Gardner Rechnung exakte Fehlerfreiheit verlangt, fordert die Willshaw Rechnung lediglich asymptotisch verschwindendes Fehlerverhältnis (3.18).

Ausgangskapazität

Informationsverluste durch Retrievalfehler werden in der Definition der Ausgangskapazität miteinbezogen. In [Hr89] wird die Ausgangskapazität für optimales a und optimale Lernregel bei inkrementeller Speicherung für fluktuierendes p in den Inhaltsmuster berechnet. Für nichtspärliche Muster wird fast genau der Zahlenwert der Musterkapazität von Amit et al reproduziert, die Retrievalfehler reduzieren also die Kapazität kaum merklich.

Auch für spärliche Muster und $a = 0$ wird für $p \to 0$ ein Wert der Ausgangskapazität angenommen, der mit dem von Tsodyks übereinstimmt. Genau wie bei Einschrittretrieval[9] wird aber dieser Wert im spärliche Limes sehr langsam angenommen, beispielsweise wird für $p = 10^{-5}$ ungefähr gerade der halben Wert erreicht.

Die Behandlung des dynamischen Problems bei inkrementeller Speicherung mit der optimalen Regel und $\mu = 0$ ergab in [HBFKS89], daß Inhaltsmuster mit $p = 0.1$ bereits mit $K = 0.143$ Bit/Synapse einen höheren Wert der Korrekturkapazität erbringen gegenüber $K = 0.092$ Bit/Synapse für $p = 0.5$.

Unsere Abschätzung durch die Fixpunktkapazität zeigt für die Hebbregel, daß prinzipiell in den Fixpunkten genau die doppelte Menge an Information steckt (und durch das Fixpunkt-Ausleseverfahren auslesbar ist), die durch Einschrittretrieval zugänglich ist, nämlich $F = 0.35$ Bit/Synapse, bzw. $F = 0.36$ Bit/Synapse. Für inkrementelle Speicherung wurde in [Pa92] die lokale Lernschranke L_A^* angegeben, eine obere Schranke der mit lokalen Lernregeln erreichbaren Lernkapazität[10]. Die Fixpunktkapazität F für inkrementelle Speicherung stimmt genau mit der lokalen Lernschranke überein, was dahingehend interpretiert werden kann, daß die Auswertung aller Fixpunkte eine völlig verlustfreie Auslesemethode darstellt.

Für die Hebbregel bei binärer Speicherung konnten wir in Abschnitt 6.3.1 mit sehr einfachen Argumenten zeigen, daß sich die Korrekturkapazität von $K = 0.17$ Bit/Synapse beim Übergang von Einschrittretrieval zu iterativem Retrieval mit der Schwellenstrategie Tk+ nicht erhöht, daß also für Mustervervollständigung die Kapazität der Fixpunkt-Auslesemethode asymptotisch nicht erreichbar ist.

7.2.2 Endliche Systemgröße

Inkrementelle Speicherung

Im Vergleich zu thermodynamischen Systemen ist bei der Beschreibung von durch Hard- oder Software realisierbaren neuronalen Netzwerken der thermodynamische Limes weitaus schlechter zu rechtfertigen. Während das Mol eines Gases 10^{23} Teilchen enthält, umfassen künstliche Netzwerke weniger als 10^8 Neuronen, ja selbst die geschätzte Zahl menschlicher Gehirnzellen liegt nicht über 10^{11}. Deshalb sind Theorien für endliche Systemgrößen bei neuronalen Netzen von sehr viel größerer Bedeutung als in der Thermodynamik.

---

[9] siehe Abbildung 4.4

[10] Die Lernkapazität ist definiert als Kanalkapazität des linken Abschnitts im Datenkanal aus Abbildung 3.4, des Lernprozesses.

Die höchsten asymptotischen Kapazitätswerte werden für spärliche Muster, genauer im spärlichen Limes erreicht. Für inkrementelle Speicherung wurde in [HBFKS89] darauf hingewiesen, daß sie erst für extrem kleines p – wie es nur für „astronomische" Systemgröße realisierbar ist – erreicht werden.

Unsere Simulationen mit der Hebbregel bei inkrementeller Speicherung benutzen $a = 0$, also nicht den nach Abschnitt 4.5.5 optimalen a Wert wie in [FL92]. Hierdurch erhalten wir auch bei optimiertem p noch kleinere Kapazitäten als in [FL92] (maximal um $K = 0.07$ Bit/Synapse, siehe Abbildung 6.3).

Binäre Speicherung
Auch binäre Speicherung wird meist, z. Bsp. in [GRS90], für ein „realistisches" p mit $p >> \ln[n]/n$ untersucht (und damit ein sehr viel niedrigerer Musterkapazitätswert erreicht). Die Forderung $p \propto \ln[n]/n$ ist aber für Anwendungen nicht prinzipiell unbrauchbar oder unrealistisch. Bereits für praktikable Systemgrößen kann man einige hundert Bit an Information in ein Inhaltsmuster kodieren.

Wie die Resultate der vorliegenden Arbeit zeigen, werden bei binärer Speicherung auch bereits bei durchaus simulierbaren Systemgrößen die asymptotischen Werte – die gegenüber den anderen Modellen neuronaler Assoziativspeicher ohnehin höher liegen – beinahe erreicht, beziehungsweise für iteratives Retrieval sogar überschritten.

Analyse und Experiment ergeben, daß Einschrittretrieval den asymptotischen Wert relativ rasch erreicht, für $n = 20000$ bereits 95 %.

Für die Korrekturkapazität des zweiten Retrievalschrittes zeigt Analyse und Experiment ein noch komplexeres Verhalten. Der asymptotische Wert wird sogar zunächst überschritten, es wird für endliches n ein Maximum durchlaufen und schließlich für sehr große n (weit außerhalb experimentell zugänglicher Systemgrößen) der Grenzwert von oben approximiert.

Iteratives Retrieval, daß bis zum Fixpunkt geführt wird, scheint das gleiche Verhalten aufzuweisen wie Retrieval bis zum zweiten Schritt. Dies läßt sich schließen, einerseits aus den experimentellen Werten, die noch über den Werten des zweiten Schrittes liegen und andererseits aus dem hier ermittelten asymptotischen Grenzwert. Auch für inkrementelle Speicherung liegt die Korrekturkapazität, die in [FL92] mit Simulationen erreicht wird, über den asymptotischen Werten aus [HBFKS89][11].

7.3 Konsequenzen für die Anwendung

Es wurde bereits in früheren Arbeiten darauf hingewiesen [WBL69, Pa80, Hr89], daß bei Speicherung spärlicher Muster mit der Hebbschen Lernregel die höchsten Speicherkapazitäten erreichbar sind.

Aus zwei Gründen ist für die technische Anwendung die binäre Speicherung zu favorisieren: Ein Gesichtspunkt ist die Speichereffizienz: bei fast unveränderten Kapazitätswerten wird zur Darstellung einer Synapse lediglich ein Bit benötigt. Ferner zeigt sie die beste Performanz für $a = 0$, gerade für spärliche Muster bedeutet $a \neq 0$ einen wesentlich erhöhten Signalfluß und Rechenaufwand im Retrieval.

Behält man für inkrementelle Speicherung das (a=0)-Modell bei, so ist bei endlicher Systemgröße die Bedingung $\mu = 0$ nicht erfüllt. Der experimentelle Vergleich bis zu Systemgrößen von $n = 20000$ ergibt, daß inkrementelle Speicherung bei solchen Systemgrößen

---

[11] Möglicherweise deutet diese Diskrepanz zur Theorie auf ein ähnliches Verhalten des inkrementellen Modells für endliche Systemgröße hin.

noch sehr viel schwächer abschneidet: in unseren Experimenten wurden nur die halben Korrekturkapazitätswerte gegenüber binärer Speicherung erzielt.

Für Autoassoziation im Modell binärer Speicherung mit der Hebbregel konnte hier gezeigt werden, daß bei geeigneter Schwellenstrategie durch iteratives Retrieval die Korrekturkapazität für endliche Systemgröße ($n = 1900$) gegenüber Einschrittretrieval um bis zu ca. 20% auf 0.19 Bit/Synapse erhöht und die Retrievalfehler gesenkt werden, die man bei optimaler Speicherbeladung erhält. Ganz generell sind die Iterationszeiten kurz (< 10).

Wir konnten eine neue Schwellenstrategie angeben, welche Zweierzyklen vermeidet und außerdem die kürzesten Iterationszeiten von allen hier untersuchten Strategien aufweist (Strategie Tk+).

Zur Realisierung des binären neuronalen Assoziativspeichermodells gibt es zwei Ansätze. Die Implementierung auf einem herkömmlichen Rechner, wie sie in unseren Simulationen verwendet wurde, werden die spärlichen Muster durch Pointer auf ihre Einskomponenten repräsentiert und als solche verarbeitet. Seit einigen Jahren wird eine Hardwarerealisierung im PAN-Projekt entwickelt. Bei den ersten Versionen PAN I und PAN II wurde eine kleine Anzahl (9 bzw. 23) von Z80 Prozessoren auf einen gemeinsamen BUS parallelgeschaltet und auf jedem Prozessor eine feste Zahl von Neuronen simuliert [PB84, Er91]. PAN III verwendet herkömmliche RAM Bausteine für die binären Synapsen. Für die neuronalen Operationen des Aufsummierens und des Schwellenvergleichs wurde eine VLSI Chip entwickelt [HPG90].

Zu einer Nutzung des binären Assoziativspeichers sind Algorithmen notwendig, die spärliche Muster möglichst konstanter Aktivierung für Anwendungen erzeugen. Für Wörter wurde in [BHP89] eine spärliche Kodierung vorgeschlagen, welche im Wort vorhandene Buchstabentrippel auf Mustereinsen abbildet. Für die autoassoziative Verarbeitung von Grauwertbildern wurde in [Ze90] gezeigt, daß durch die Reduktion der Bildentropie, wie sie in Low-level vision oft angewendet wird, eine spärliche Kodierung erreicht werden kann.

Zum generellen Problem der Erzeugung ähnlichkeitserhaltender Kodierungen entstanden Arbeiten in unserer Gruppe [Sm92, PSS93].

7.4 Konsequenzen für die Hirnforschung

Bei einer Modellierung des biologischen Nervensystems, die nicht zwangsläufig mit der Beschäftigung mit neuronalen Netzmodellen gleichgesetzt werden sollte, gibt es zwei extreme Vorgehensweisen: Entweder wird das Neuron sehr detailliert modelliert, wegen der Komplexität des Einzelneurons kann das Modell dann aber nur wenig Neuronen enthalten, siehe Abschnitt 2.5. Das andere Extrem bildet die Modellierung durch den hier betrachteten Assoziativspeicher. Es ist kaum ein Modell vorstellbar, daß eine stärkere Abstraktionen des biologischen Nervensystems darstellt, als die hier untersuchten Speichermodelle. Durch die radikale Vereinfachung der Details läßt sich das Verhalten einer sehr große Anzahl von Modellneuronen untersuchen. Diese stark abstrahierende Vorgehensweise der Modellierung geht von der Hypothese aus, daß das kollektive Verhalten der Neurone, welches nach der allgemein akzeptierten Auffassung für die charakteristischen Eigenschaften des Nervensystem verantwortlich ist, robust ist gegenüber der Variation der mikroskopischen Details. Eine solche Hypothese wird gestützt durch Erfolge der statistischen Mechanik im Verständnis thermodynamischer Phänomene durch mikroskopische Ansätze, z. Bsp. bei Phasenübergängen.

Viele unserer Modelleigenschaften, so die lokalen Lernregeln und die Anwesenheit von Rückkopplung stellen grundlegende Merkmale des Nervensystems da. Auch ist die eine der beiden effektiven Schwellenstrategien, die hier untersucht wurde, die Strategie CA, biologisch plausibel.

Zunächst erscheinen binäre Synapsen biologisch weniger plausibel. Es sollte allerdings, wegen der begrenzten Genauigkeit der biologischen Synapsen, auch nicht von einer allzugroßen Tiefe der biologischen synaptischen Gewichte ausgegangen werden.

Das Ergebnis, daß in unserem Modell Mustervervollständigung von spärlichen Mustern mit kurzen Iterationszeiten maximaler Korrekturkapazität möglich ist, kann zumindest als Hinweis verstanden werden, daß auch im Nervensystem, wo man ebenfalls spärliche Aktivitätsmuster und kurze Iterationsdauer beobachtet, elementare Assoziationsleistungen wie Mustervervollständigung eine wichtige Rolle bei der Informationsverarbeitung spielen.

7.5 Resumé

Zu Beginn der Erstellung der vorliegenden Arbeit zum iterativen Retrieval stand ein eingehendes Studium der thermodynamischen Methoden und Arbeiten, die sich mit (rückgekoppelten) Assoziativspeichermodellen beschäftigen. Durch die thermodynamische Analyse war bereits das asymptotische Verhalten der Musterkapazität bei inkrementeller Speicherung sowohl von nichtspärlichen Lernmustern, als auch im spärlichen Limes bekannt.

Der für spärliche Muster wichtigen Frage nach der tatsächlichen Informationskapazität wurde in dieser Literatur nur wenig Aufmerksamkeit geschenkt[12]. Daher wurde hier der informationstheoretische Aspekt und die daraus resultierenden Kapazitätsdefinitionen von Palm [Pa91] nocheinmal ausführlich dargestellt.

Über binäre Speicherung lagen durch die Arbeiten mit elementar statistischen Methoden fast nur Ergebnisse für Einschrittretrieval vor. Die Kapazitätswerte für spärliche Inhaltsmuster waren vielversprechend und gaben mit den Auschlag zum Hauptgegenstand dieser Arbeit: eine Untersuchung von iterativem Retrieval für binäre Speicherung mit der Hebbregel. Speziell für diese Modellvariante wurden Schwellenstrategien entwickelt und mit den Simulationsexperimenten begonnen. Die Experimente zeigten, daß bereits bei Systemgrößen, die gut in Simulation und Anwendung realisierbar sind, die asymptotischen Kapazitätswerte für Einschrittretrieval bereits im ersten Retrievalschritt beinahe erreicht und in weiteren Retrievalschritten oft sogar überschritten werden.

Unsere Berechnung der Fixpunktkapazität durch die kombinatorisch statistische Analyse des Fixpunkt-Ausleseverfahrens stellt ein neues Verfahren dar, die Information über die Inhaltsmuster abzuschätzen, die aus den Fixpunkten extrahiert werden kann. Die Zahlenwerte der Fixpunktkapazität, als Abschätzung für die asymptotische Korrekturkapazität von Fixpunktretrieval interpretiert, deuten auf eine Verbesserung durch Iteration hin; sie sind genau doppelt so hoch wie die Korrekturkapazität des Einschrittretrieval.

Hingegen zeigt unsere asymptotische Analyse des iterativem Retrieval, daß die Korrekturkapazität mit derjenigen des Einschrittretrieval zusammenfällt, daß also mit iterativem Retrieval asymptotisch keine Verbesserung gegenüber Einschrittretrieval erzielt wird. Die Fixpunktkapazität als eine obere Schranke ist durch Auslesemethoden, die zur Mustervervollständigung einsetzbar sind, nicht erreichbar. Diese Aussage der asymptotischen Theorie stand jedoch im Wiederspruch zu unseren Experimenten. In Simulationen beobachteten wir eine deutliche Erhöhung der Korrekturkapazitätswerte durch die Iteration des Retrieval; die experimentellen Werte lagen über dem asymptotischen Theoriewert. Eine Klärung dieses Sachverhalts versprachen wir uns durch eine theoretische Beschreibung des endlichen Systems.

Während die thermodynamischen Methoden ausschließlich asymptotische Ergebnisse liefern, gab es bereits Arbeiten, die mit elementar statistischen Methoden das Retrieval im

---

[12] Eine Ausnahme bilden die Arbeiten von Horner et al [HBFKS89, Hr89, FL92].

endlichen System untersuchen[13]. Für extreme Spärlichkeit, die bisher noch nicht behandelt war, konnte in der vorliegenden Arbeit für den zweiten Retrievalschritt eine Analyse gerechnet werden. Die Rechnung ergibt in Übereinstimmung mit unseren Experimenten, daß bei binärer Speicherung und iterativem Retrieval für endliches System der asymptotische Wert überschritten wird und sehr weit draußen von oben approximiert wird.

Insgesamt konnte durch diese Arbeit ein Verständnis für das Verhalten der für die Vervollständigung von Zufallsmustern effektivsten und zudem auch anwendbaren Modellvariante mit iterativem Retrieval gefunden werden. Es bleiben freilich noch Fragen offen, so ist die endliche Theorie des zweiten Schrittes für kleine Systemgrößen zu optimistisch und es gibt noch keine endliche Theorie für weitere Retrievalschritte. Auch wäre eine Analyse des inkrementellen Modells mit optimaler Lernregel im endlichen Systems noch interessant. Für die Anwendung solcher neuronaler Speichermodelle stehen aber eher andere Fragen im Vordergrund, so das Verhalten des Modells für korrellierte Muster und das spärliche Kodierungsproblem.

---

[13] vgl. die Arbeit von Gibson und Robinson [GR91].

Literaturverzeichnis

[Ai88] Amari, S. I.: Statistical neurodynamics of associative memory. Neural Networks **1** (1988) 63 - 73

[Ai89] Amari, S. I.: Mathematical foundations of neurocomputing. Technical report METR 89 - 06 University of Tokyo (1989)

[At89] Amit, D.J.: Modelling brain function. Cambridge University Press (1989)

[AGS87] Amit, D.J., Gutfreund, H., Sompolinsky H.: Statistical mechanics of neural networks near saturation. Ann. Phys. **173** (1987) 30 - 67

[An68] Anderson, J. A.: A memory storage model utilizing spatial correlation functions. Kybernetik **5** (1968) 113 - 119

[An72] Anderson, J. A.: A simple neural network generating an interactive memory. Math. Biosci. **14** (1972) 197 - 220

[BHP89] Bentz, J. B., Hagstroem, M., Palm, G.: Information storage and effective data retrieval in sparse matrices. Neural Network **2** (1989) 289 - 293

[BSB81] Barto, A.G., Sutton, R.S., Brouwer, P.S.: Associative search network; a reinforcement learning associative memory. Biol. Cybern. **40** (1981) 201 - 211

[BDS89] Buhmann, J., Divko, R., Schulten, K.: Associative memory with high information content. Phys. Rev. A **39** 1989) 2689 - 2692

[Bu88] Buhmann, J.: Neuronale Netzwerke als assoziativer Speicher und als Systeme zur Mustererkennung. Dissertation TU München (1988)

[Co65] Cover, T. M.: Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition. IEEE Transactions on electronic computers (1965) 326 - 334

[CP43] Mc Culloch, W. S., Pitts, W.: A logical calculus of the ideas immanent in neural activity. Bul. of Math. Biophys. **5** (1943)

[DHS91] Domany, E., van Hemmen, J.L., Schulten, K: Models of neural network. Springer (1991)

[Er91] Erb, M. K.: Simulation neuronaler Netze: Stabilität, Plastizität und Konnektivität. Dissertation Universität Tübingen (1992)

[FK88] Fontanari, J. F., Köberle, R.: Information processing in synchronous neural networks. J. Phys. France **49** (1988) 13 - 23

[FL92] Forrest, B. M., Loettgers, A.: Neural networks with low activity levels. Z. Phys. B (1992)

[FM88] Faris, W. F., Maier, R. S.: Probabilistic analysis of a learning matrix. Adv. Appl. Prob. **20** (1988) 695 - 705

[Ga88] Gardner, E.: The space of interactions in neural network models. J. Phys. A **21** (1988) 257 - 270

[GDM87] Gardner, E., Derrida, B., Mottishaw, P.: Zero temperature parallel dynamics for infinite range spin glasses and neural networks. J. Phys. France **48** (1987) 741 - 755

[GGY89] Gardner, E., Gutfreund, H., Yekutieli, I.: The phase space of interactions in neural networks with definite symmetry. J. Phys. A **22** (1989) 1995 - 2008

[G-M76] Gardner-Medwin, A.R.: The recall of events through the learning of associations between their parts. Proc. R. Soc.. Lond. B. **194** (1976) 375 - 402

[GR91] Gibson, W.G., Robinson, J.: Statistical analysis of the dynamics of a sparse associative memory. Neural Networks (1991)

[GRS90] Golomb, D., Rubin, N., Sompolinsky, H.: Willshaw model: associative memory with sparse coding and low firing rates. Phys. Rev. A **41** (1990) 1843 - 1854

[GS90] Gutfreund, H., Stein, Y.: Capacity of neural networks with discrete synaptic couplings. J. Phys A **23** (1990) 2613 - 2630

[Gu90] Gutfreund, H.: From statistical mechanics to neural networks and back. Physica A **163** (1990) 373 - 385

[HBFKS89] Horner, H., Bormann, D., Frick, M., Kinzelbach, H, Schmidt, A.: Transients and basins of attraction in neural network models. Z. Phys. B **76** (1989) 381 - 398

[Hd82] Hopfield, J. J.: Neural networks and physical systems with emergent collective computational abilities. Proc. Natl. Sci. **79** (1982) 2554 - 2558

[He49] Hebb, D. O.: The organization of behaviour. Wiley, New York (1949)

[HH52] Hodgkin, A. L., Huxley, A. F.: A quantitative description of membrane current and its application to conduction and excitation in nerve. J. Physiol. (Lond.) **117** (1952) 500 - 544

[HKP91] Hertz, J., Krogh, A. , Palmer, R. G.: Introduction to the theory of neural computation. Addison Wesley (1991)

[HO90/1] Henkel, R. D., Opper, M.: Distribution of internal fields and dynamics of neural networks. Europhys. Lett. (1990)

[HO90/2] Henkel, R. D., Opper, M.: Parallel dynamics of neural network with pseudoinverse coupling matrix. Europhys. Lett. (1990)

[HPG90] Huch, M., Pöchmüller, W., Glesner, M.: Bacchus: a VLSI architecture for a large binary associative memory. Proc. of INNC Paris (1990)

[HSKH88] Herz, A., Sulzer, B., Kühn, R., van Hemmen, J.L.: The Hebb rule: storing static and dynamic objects in an associative neural network. Europhys. Lett. **7** (1988) 663 - 669

[Hr89] Horner, H.: Neural networks with low levels of activity: Ising vs. McCulloch-Pitts neurons. Z. Phys. B **75** (1989) 133 - 136

[Ko77] Kohonen T.: Associative memory. Springer, Berlin (1977)

[Ko89] Kohonen T.: Self-organization and associative memory. Springer, Berlin (1989)

[KS78] Kirkpatrick, S., Sherrington, D.: Infinite-ranged models of spin-glasses. Phys. Rev. B **17** (1978) 4384 - 4403

[Kü91] Kühn, R.: Zur statistischen Physik neuronaler Netze. Habilitationsschrift (1991) Universität Heidelberg

[Li74] Little, W.A.: The existence of persistent states in the brain. Math. Biosci. **19** (1974) 101 - 120

[LWB70] Longuett-Higgins, H.C., Willshaw, D.J., Buneman, O.P.: Theories of associative recall. Q. Rev. Biophys. **3** (1970) 223 - 244

[Ma91] Marr, D.: Simple memory: a theory for archicortex Philos. Trans. R. Soc. London Ser. B **262** (1971) 23 - 81

[Na91] Nadal, J.-P.: Associative memory: on the (puzzling) sparse coding limit. J.Phys A **24** (1991) 1093 - 1101

[NT90] Nadal, J.-P., Toulouse, G.: Information storage in sparsely coded memory nets. Network **1** (1990) 61 - 74

[OKKK89] Opper, M., Kleinz, J., Köhler, H., Kinzel, W.: Basins of attraction near saturation near the critical storage capacity for neural networks with constant stabilities. J. Phys. A **22** (1989) L407 - L411

[Pa80] Palm, G.: On associative memory. Biol. Cybern. **36** (1980) 19 - 31

[Pa81] Palm, G.: On the storage capacity of an associative memory with randomly distributed storage elements. Biol. Cybern. **39** (1981) 125 - 127

[Pa82] Palm, G.: Neural assemblies. Springer, Berlin (1982)

[Pa87] Palm, G.: Computing with neural networks. Science **235** (1987) 1227 - 1228

[Pa88] G. Palm, On the asymptotic storage capacity on neural networks, *Neural Computers, Ed. R. Eckmiller, Chr. v.d. Malsburg, Springer 1988*

[Pa90] Palm, G.: Local learning rules and sparse coding in neural networks. Advanced Neural Computers Ed: R. Eckmiller Elsevier Science Publishers B.V. (North Holland) (1990) 145 - 150

[Pa91] Palm, G.: Memory capacities of local rules for synaptic modification. Concepts in Neuroscience **2** (1991) 97-128

[Pa92] Palm, G.: On the information storage capacity of local learning rules. Neural Computation **4** (1992) 703 - 711

[PB84] Palm, G., Bonhoeffer, T.: Parallel processing for associative and neural networks. Biol. Cybern. **40** (1984) 201 - 204

[PDT86] Personnaz, L., Dreyfus, G., Toulouse, G.: A biologically constrained learning mechanism in networks of formal neurons. J. Stat. Phys. **43** (1986) 411 - 422

[Pe88] Peretto, P.: On learning rules and memory storage abilities. J. Phys. France **49** (1988) 711 - 726

[PGD86] Personnaz, L., Guyon, I., Dreyfus, G.: Collective computational properties of neural networks: new learning mechanisms. Phys. Rev. A **34** (1986) 4217 - 4228

[PP91] Palm, G., Palm, M.: Parallel associative networks: the PAN-system and the BACCHUS-chip. Proc. of the 2nd International Conference on Microelectronics for neural networks. Eds: U. Ramacher, U. Rückert, J. A. Nossek Kyrill & Method Verlag

[PS91] Palm, G., Sommer, F. T.: Information and pattern capacities in neurla associative memories with feedback for sparse memory patterns. Neural Network Dynamics Eds: J.G. Taylor, E.R. Caianiello, R.M.J.Cotterill, J.W.Clark, Springer (1991) 3 - 18

[PS92] Palm, G., Sommer, F. T.: Information capacity in recurrent McCulloch - Pitts networks with sparsely coded memory states. Network **3** (1992) 1 - 10

[PS93] Palm, G., Sommer, F. T.: Associative data storage and retrieval in neural networks. In: Models of Neural Networks Eds: E. Domany, J.L. van Hemmen, K. Schulten, Springer (1993)

[PSS93] Palm, G., Schwenker, F., Sommer, F.T.: Associative memory and sparse similarity perserving codes. In: From Statistics to Neural Networks: Theory and Pattern Recognition Applications, Ed: Cherkassky, V., Springer NATO ASI Series F, Springer, New York (1993)

[PSS93] Palm, G., Stellmann, U., Schwenker, F.: Analysis, representation and simiarity preserving coding of large highdimensional data sets. WINA Bericht 1993

[P-VA89] Perez Vicente, C. J., Amit, D.J.: Optimized network for sparsely coded patterns. J. Phys A **22** (1989) 559 - 569

[PZ91] Patrick, A. E., Zagrebnov, V. A.: On the parallel dynamics of neural network models. J. Stat. Phys. **63** (1991) 59 - 71

[Ro62] Rosenblatt, F.: Principles of neurodynamics. Spartan Books, New York (1962)

[SSP93] Schwenker, F., Sommer, F. T., Palm, G.: Iterative retrieval of sparsely coded patterns in associative memory. Neuronet'93 Prague (1993)

[Sm92] Stellmann, U.: Dissertation Universität Ulm (1992)

[St61] Steinbuch, K.: Die Lernmatrix. Kybernetik **1** (1961) 36

[TF88] Tsodyks, M. V., Feigelman, M. V.: The enhanced storage capacity in neural networks with low activity level. Europhys. Lett. **6** (1988) 101 - 105

[Ts88] Tsodyks, M. V.: Associative memory in asymmetric diluted network with low level of activity. Europhys. Lett. **7** (1988) 203 - 208

[Ut56] Uttley, A. M.: Automata studies. Shannon, C. E., Mc Carthy, J. (eds) Princeton Univ. Press (1956) 237 - 252

[vH87] van Hemmen, J.L.: Nonlinear networks near saturation. Phys. Rev. A **36** (1987) 1959 - 1962

[WBL69] Willshaw, D.J., Buneman, O.P., Longuet-Higgins, H.C.: Nonholographic associative memory. Nature (London) **222** (1969) 960 - 962

[WD90] Willshaw, D.J., Dayan, P.: Optimal plasticity from matrix memories: what goes up must come down. Neural Computation **2** (1990) 85 - 93

[WD91] Willshaw, D.J., Dayan, P.: Optimizing synaptic learning rules in linear associative memories. Biol. Cybern. (1991) 253 - 265

[Ze90] Zetsche, C.: Sparse coding: the link between low level vision and associative memory. in: Parallel Processing in Neural Systems and Computers, Eds.: Eckmiller, R., Hartmann, G., Hausske, G. Elsevier (1990)

Anhang A

Binäre Hebbregel

Besitzt das Eingangsmuster im Retrieval l Einsen, dann werden genau l Elemente einer Spalte j der Synapsenmatrix zum dendritischen Potential aufsummiert. Falls die Matrixelemente statistisch unabhängig sind mit $p[\bar{c}_{ij} = 1] = 1 - p_0$, ist das dendritische Potential eine binomialverteilte Zufallsvariable mit dem Mittelwert $l(1 - p_0)$.

Es ist jedoch durch die Erzeugung der Matrixelemente im Lernprozess auch bei zufallsgenerierten Lernmustern, wie wir sie hier verwenden, keine strenge Unabhängigkeit der Matrixelemente gegeben.

Proposition 1
In einer Spalte der Hebbmatrix $\bar{c}$, die der binäre Speicherprozess aus einer Menge von M mit $Z(p)$ erzeugten Zufallsmustern mit den Aktivierungen p und q generiert, gilt für l Einträge:

$$\begin{aligned} p[\cap_{i=1}^{l}(\bar{c}_{ij}=1)] &= 1+\sum_{i=1}^{l}\binom{l}{i}(-1)^i\left[1-q\left\{1-(1-p)^i\right\}\right]^M \\ &\approx 1+\sum_{i=1}^{l}\binom{l}{i}(-1)^i\exp\left[-Mq\left\{1-(1-p)^i\right\}\right] \end{aligned} \tag{A.1}$$

Insbesondere folgt hieraus mit der Definition $V := Mpq$ für ein Matrixelement

$$p_0 := 1 - p[\bar{c}_{ij}=1] = (1-pq)^M \approx e^{-V} \tag{A.2}$$

und für zwei Matrixelemente

$$\mathrm{Cov}[\bar{c}_{1j}\bar{c}_{2j}] = (1-q(2p-p^2))^M - (1-pq)^{2M} \approx Vpp_0^2. \tag{A.3}$$

Beweis
Seien wieder x_i^k, y_j^k die Lernmusterpaare mit dem Musterindex $1 \leq k \leq M$ und i und j die Komponentenindizes eines Adressmusters, bzw. eines Inhaltsmusters. Für ein Matrixelement ist

$$p[\bar{c}_{1j}=1] = p[\exists k : x_1^k y_j^k] = 1 - p[\forall k : x_1^k y_j^k = 0] = 1-(1-pq)^M.$$

Für zwei Matrixelemente gilt

$$\begin{aligned} p[(\bar{c}_{1j}=1)\cap(\bar{c}_{2j}=1)] &= p[(\exists k : x_1^k x_2^k y_j^k)\cup(\exists l,m : x_1^l(1-x_2^l)y_j^l(1-x_1^m)x_2^m y_j^m)] \\ &= 1-(p[E_1]+p[E_2]-p[E_1\cap E_2]), \end{aligned}$$

mit den Ereignissen

$$\begin{aligned} E_1 &:= [\forall k : x_1^k x_2^k y_j^k = 0 \cap x_1^k(1 - x_2^k) y_j^k = 0], \\ E_2 &:= [\forall k : x_1^k x_2^k y_j^k = 0 \cap (1 - x_1^k) x_2^k y_j^k = 0]. \end{aligned}$$

Es ist $p[E_1] = p[E_2] = (1 - pq)^M$ sowie $p[E_1 \cap E_2] = (1 - q(2p - p^2))^M$ und wir erhalten

$$p[(\bar{c}_{1j} = 1) \cap (\bar{c}_{2j} = 1)] = 1 - \binom{2}{1}(1 - pq)^M + (1 - q(2p - p^2))^M.$$

Die Kovarianz ergibt sich aus den Berechungen für $l = 1$ und $l = 2$:

$$\begin{aligned} \mathrm{Cov}[\bar{c}_{1j}, \bar{c}_{2j}] &= p[(\bar{c}_{1j} = 1) \cap (\bar{c}_{2j} = 1)] - p[\bar{c}_{1j} = 1]^2 \\ &= (1 - 2qp + qp^2)^M - (1 - 2pq + p^2 q^2)^M \\ &\approx e^{-2pqM}(e^{Mp^2 q} - e^{Mp^2 q^2}) \approx p_0^2 M p^2 q(1 - q) \approx p_0^2 M p^2 q \end{aligned}$$

Für $l > 2$ kann die Berechnung von $p[\cap_{i=1}^l (\bar{c}_{ij} = 1)]$ nach dem gleichen Schema fortgesetzt werden. ◇

Proposition 2

In einer Spalte der Hebbmatrix $\bar{c}$, die wie in Proposition 1 generiert wurde, jedoch für Autoassoziation und aus Mustern die mit $Z(k)$ erzeugt wurden, gilt für l Einträge:

$$p[\cap_{i=1}^l (\bar{c}_{ij} = 1)] = 1 + \sum_{i=1}^{l} \binom{l}{i} (-1)^i \left[\frac{T(n,k,i)k}{n - i} + \frac{n - k}{n}\right]^M \tag{A.4}$$

mit

$$T(n,k,i) := \prod_{h=0}^{i-1} \frac{n - k - h}{n - h}. \tag{A.5}$$

Insbesondere folgt hieraus für ein Matrixelement

$$p_0 = \left[1 - \frac{kk'}{nn'}\right]^M < \left[1 - \left(\frac{k}{n}\right)^2\right]^M \tag{A.6}$$

und für zwei Matrixelemente

$$\mathrm{Cov}[\bar{c}_{1j} \bar{c}_{2j}] \approx \frac{kk'}{nn'} \left[\frac{k - 2}{n - 2} - \frac{kk'}{nn'}\right] M \exp\left[-2M \frac{kk'}{nn'}\right] \tag{A.7}$$

mit $n' = n - 1$ und $k' = k - 1$.

Für große Systemgröße n und $k/n = p = q$ gehen diese Gleichungen in diejenigen aus Proposition 1 über.

Beweis

Siehe [Pa80], für große n wegen Proposition 1. ◇

Die Tatsache, daß generell die statistische Unabhängigkeit der Spaltenelemente nicht erfüllt ist, erkennt man am Kovarianzterm, der im allgemeinen nicht verschwindet. Die folgende Proposition gibt die Bedingungen an, unter welchen die Unabhängigkeitsannahme aus Abschnitt 4.4.2 gerechtfertigt ist.

Proposition 3
In einer Spalte der Hebbmatrix, die wie in Proposition 1 erzeugt wurde, gilt

$$\frac{p[\cap_{i=1}^{l}(\bar{c}_{ij}=1)]}{p[\bar{c}_{ij}=1]^{l}} \to 1 \text{ für } n \to \infty, \tag{A.8}$$

so lange einerseits $pl^2 \to 0$ aber andererseits $V = Mpq > 0$ für $n \to \infty$.

Beweis:
Es gilt

$$\begin{aligned} p[\bar{c}_{ij}=1] &\leq p[\bar{c}_{lj}=1 | \cap_{i=1}^{l-1}(\bar{c}_{ij}=1)] \leq p[\bar{c}_{lj}=1 | \sum_{k=1}^{M} y_j^k \geq l-1] \\ &= 1-(1-p)^{l-1}(1-pq)^{M-l+1}. \end{aligned}$$

Damit erhalten wir die oberen Abschätzungen

$$\begin{aligned} 0 &\leq \log\left[\frac{p[\cap_{i=1}^{l}(\bar{c}_{ij}=1)]}{p[\bar{c}_{ij}=1]^{l}}\right] \leq \sum_{i=0}^{l-1} \log\left[\frac{1-(1-p)^{i}(1-pq)^{M-i}}{1-(1-pq)^{M}}\right] \\ &= \sum_{i=0}^{l-1} \log\left[\frac{1-(\frac{1-p}{1-pq})^{i}p_0}{1-p_0}\right] \\ &\leq \sum_{i=0}^{l-1} \log\left[\frac{1-(1-ip)p_0}{1-p_0}\right], \text{ wegen } (\frac{1-p}{1-pq})^{i} \geq (1-p)^{i} \geq 1-ip, \\ &\leq \sum_{i=0}^{l-1} ip\frac{p_0}{1-p_0}, \text{ wegen } \log[1+x] \leq x, \\ &\leq \frac{p \cdot p_0}{1-p_0} \cdot \frac{l^2}{2} \to 0 \text{ für } p \cdot l^2 \to 0 \text{ und falls } p_0 \not\to 1, \text{ also } V > 0. \end{aligned}$$

◇

Für Einträge in einer Zeile der Matrix gelten die selben Gleichungen, lediglich mit vertauschtem p und q. Matrixelemente in verschiedenen Spalten *und* Zeilen sind statistisch unabhängig.

Die folgende Proposition gibt Abschätzungen an, welche die statistischen Abhängigkeiten mehrerer Matrixelemente einer Spalte mitberücksichtigen.

Proposition 4
In einer Spalte der Hebbmatrix $\bar{c}$, die der binäre Speicherprozess aus einer Menge von M Zufallsmustern generiert, gilt für $l+j$ Einträge:

$$p[\cap_{i=1}^{l}(\bar{c}_{ij}=1) \cap_{i=l+1}^{l+j}(\bar{c}_{ij}=0)] = p[A]\, p[B], \tag{A.9}$$

wobei

i) fr Adress- und Inhaltsmuster, die jeweils mit mit der Aktivierung p (Autoassoziation) durch $Z(p)$ erzeugt wurden, gilt

$$p[A] = 1 + \sum_{h=1}^{l} \binom{l}{h} (-1)^h \left[1 - p(1-p)^j \left\{1-(1-p)^h\right\}\right]^{M-1} \tag{A.10}$$

$$p[B] = \left[1 - p\left\{1-(1-p)^j\right\}\right]^{M-l} \tag{A.11}$$

ii) für Adress- und Inhaltsmuster, die jeweils mit mit der Aktivität k durch $Z(k)$ erzeugt wurden (Autoassoziation), gilt

$$\begin{aligned} p[A] &= 1+\sum_{h=1}^{l}\binom{l}{h}(-1)^h\left[1-\left(\frac{k\,T(n,k,j)}{n-j}-\frac{k\,T(n,k,j+h)}{n-j-h}\right)\right]^{M-1} && \text{(A.12)} \\ p[B] &= \left[1-\frac{k}{n}\left\{1-T(n-1,k-1,j)\right\}\right]^{M-l} && \text{(A.13)} \end{aligned}$$

Beweis:
Ohne Beschränkung der Allgemeinheit betrachten wir Spalte 1 der Speichermatrix:

$$\begin{aligned} & p[\forall i \in \{1,2,...,l\} : \bar{c}_{i1}=1, \forall i \in \{l+1,2,...,l+j\} : \bar{c}_{i1}=0] \\ &\quad = p[\exists X := \{z^1,...,z^l\} \subset \mathcal{S} : \forall z^i : \left(z^i_i z^i_1=1\right)\cap\left(\forall p \in \{l+1,2,...,l+j\} : z^i_p z^i_1=0\right)] \\ &\quad \times \; p[\forall z \notin \mathcal{S} : \forall p \in \{l+1,2,...,l+j\} : z^i_p z^i_1=0] =: p[A]\,p[B] \qquad \text{(A.14)} \end{aligned}$$

Für den Faktor $p[B]$ erhält man (A.11) und (A.13) analog der Vorgehensweise in den Propositionen 1 und 2.

Im Faktor $p[A]$ von Gleichung (A.14) können wir die Bedingung, die das Muster $z^i \in X$ erfüllen muß, umschreiben zu

$$(z_1=1)\cap(z_i=1)\cap(\forall p \in \{l+1,2,...,l+j\} : z_p=0). \qquad \text{(A.15)}$$

Die zu (A.15) komplementäre Bedingung nennen wir *Ausfallbedingung an der Stelle i*:

$$(z_0=0)\cup(z_i=0)\cup(\exists p \in \{l+1,2,...,l+j\} : z_p=1).$$

Wird sie von allen Mustern in $\mathcal{S}$ erfüllt, dann kann es die Menge $X \subset \mathcal{S}$ nicht geben.

Mit der Ausfallsbedingung an h verschiedenen Stellen

$$(z_0=0)\cup(z_1=0, z_2=0,...,z_h=0)\cup(\exists p \in \{l+1,...,l+j\} z_p=1) =: \alpha\cup\beta_h\cup\gamma$$

erhalten wir

$$p[A] = 1+\sum_{h=1}^{l}\binom{l}{h}(-1)^h\, p[\forall z \in \mathcal{S} : \alpha\cup\beta_h\cup\gamma]. \qquad \text{(A.16)}$$

Die Ausfallsbedingung in der Wahrscheinlichkeit in Gleichung (A.16) läßt sich mittels elementarer Mengentheorie umformen zu

$$\alpha\cup\beta_h\cup\gamma = \left\{\left[\alpha\cap\gamma^C\right]\setminus\left[\alpha\cap\gamma^C\cap\beta_h\right]\right\}^C.$$

Mit γ^C bezeichnen wir die zu γ komplementäre Bedingung.

Für den Zufallsprozess $Z(p)$ erhalten wir hieraus die Wahrscheinlichkeit in Gleichung (A.16) als

$$p[\forall z \in \mathcal{S} : \alpha\cup\beta_h\cup\gamma] = \left[1-p(1-p)^j\left\{1-(1-p)^h\right\}\right]^{M-1}$$

und somit direkt Gleichung (A.10).

Analog erhalten wir für den Zufallsprozess $Z(k)$ die Wahrscheinlichkeit in Gleichung (A.16) zu

$$p[\forall z \in \mathcal{S} : \alpha\cup\beta_h\cup\gamma] = \left[1-\left(\frac{k}{n-j}T(n,k,j)-\frac{k}{n-j-h}T(n,k,j+h)\right)\right]^{M-1},$$

also direkt Gleichung (A.12). $\diamondsuit$

Anhang B

Inkrementelle Hebbregel

Proposition 1
Für das Integral über die Gaussverteilung $G[x] := (1/\sqrt{2\pi})\int_{-\infty}^{x} e^{-x^2/2}dx$ gelten die Abschätzungen:

$$(2\pi t^2)^{-1/2}e^{-t^2/2}(1-t^2) \leq G(-t) = 1 - G(t) \leq (2\pi t^2)^{-1/2}e^{-t^2/2}.$$

Beweis:
Wegen $x^2 = t^2 + (x-t)^2 + 2t(x-t)$, gilt

$$\int_t^\infty e^{-x^2/2}dx = e^{-t^2/2}\int_0^\infty e^{-x^2/2}e^{-xt}dx$$

Damit und mit der Abschätzung $e^{-x^2/2} \leq 1$ erhalten wir direkt die zweite Ungleichung weil $\int_0^\infty e^{-xt}dx = 1/t$. Die erste Ungleichung erhält man nach partieller Integration wegen $\int_0^\infty xe^{-xt}dx = 1/t$. $\diamondsuit$

Bei der Analyse des Fixpunkt-Ausleseverfahren bei inkrementeller Speicherung wird die folgende Proposition ausgenutzt:

Proposition 2
Mit den Definitionen aus Abschnitt 5.5 gilt für die inkrementelle Hebbregel:

$$Q_{ss}^{on} \propto \exp\left[\frac{-\theta^2 n^2}{4M}\right] \quad \text{für } n \to \infty \tag{B.1}$$

Beweis:
Wieder sei das abgerufene Inhaltsmuster von der Form $x = (1, 1, ..., 0, 0, ...)$, so daß wir lediglich den linken oberen k^2 Block der Matrix c_{ij} betrachten müssen.
i) Es gilt die obere Abschätzung:

$$Q_{ss}^{on} \leq p[\underbrace{\sum_{i=1,j\leq i}^{k,k} c_{ij}}_{=:s} \geq \frac{k\Theta}{2}] = G\left(-\frac{n\theta}{\sqrt{2m}}\right) \leq \frac{\sqrt{2m}}{n\sqrt{2\pi}}\exp\left[-\frac{n^2\theta^2}{4M}\right]$$

da $E(s) = k^2Mp^2/2$, $\sigma^2(s) = E(s)$ und $\Theta = kMp^2 + k\theta$, siehe Proposition 1.

ii) Sei $\{N_{ij} : i = 1, ..., k/l\}$ eine Partition jeder Spalte j mit $|N_{ij}| = l$. T bezeichnet die kleinste Menge von N_{ij} welche das linke untere Dreieck des k^2 Blocks vollständig überdeckt. Es gilt nun die Abschätzung nach unten:

$$Q_{ss}^{on} \;\geq\; p[\underbrace{\sum_{h \in N_{ij}} c_{hj}}_{=:s'} \geq \frac{\Theta l}{k} \;\;\forall\;\; N_{ij} \in T] \;=\; p\left[s' \geq \frac{\Theta l}{k}\right]^{(1+2+...+k/l)l} \tag{B.2}$$

$$= \;\; G\left(-\frac{n\theta\sqrt{l}}{k\sqrt{M}}\right)^{\frac{k^2}{2l}} \approx \left(\frac{k\sqrt{M}}{\sqrt{2\pi l}n\theta}\right)^{\frac{k^2}{2l}} \exp\left[-\frac{n^2\theta^2}{4M}\right] \tag{B.3}$$

Die letzte Abschätzung folgt aus Proposition 1 dieses Anhangs.
$\diamondsuit$

Anhang C

Fixpunkt-Ausleseverfahren

In diesem Teil skizzieren wir die Erweiterung der Rechenmethode aus Kapitel 5 auf Inhaltsmuster und Eingangsmuster, bei denen die Aktivierung p vorgegeben ist. Die Anzahl der aktiven Neurone in einem Musters ist dann, wie in Abschnitt 3.4 diskutiert, nicht in allen Inhaltsmustern gleich, sondern fluktuiert: $p[x \in \mathcal{S} \mid |x| = l] = Mp^l(1-p)^{n-l}$.

In der Analyse müssen also Muster mit unterschiedlicher Anzahl l von aktiven Neuronen miteinbezogen werden. Für die bedingte Information durch Auslesefehler muß anstelle von (5.5) aus Abschnitt 5.4.1 angesetzt werden:

$$\begin{aligned} I_{um} &= \sum_{x\in\{0,1\}^n} p[x\notin\tilde{\mathcal{S}}]\, I\left(p[x\in\mathcal{S}\,|\,x\notin\tilde{\mathcal{S}}]\right) \\ &= \sum_{l=0}^{n}\binom{n}{k} p\Big[x\notin\tilde{\mathcal{S}}\cap x\in\mathcal{S} \mid |x|=l\Big]\ \mathrm{ld}\left[\frac{p\Big[x\notin\tilde{\mathcal{S}}\mid |x|=l\Big]}{p\Big[x\notin\tilde{\mathcal{S}}\cap x\in\mathcal{S}\mid |x|=l\Big]}\right]. \end{aligned}$$

An die Stelle der Fehlerwahrscheinlichkeit p_{um} treten nun Fehlerwahrscheinlichkeiten für jedes l: $p_{um}^l := p\Big[x \notin \tilde{\mathcal{S}} \mid x \in \mathcal{S} \cap |x| = l\Big]$.

Das Hifi-Qualitätskriterium kann wieder analog zu Bedingung (5.6) in eine Bedingung an diese Fehlerwahrscheinlichkeiten übersetzt werden.

$$\frac{Mk\,\mathrm{ld}[n]}{n^2}\sum_{l=0}^{n}\binom{n}{l}p^l(1-p)^{n-l}p_{um}^l \to 0 \quad \text{für} \quad n\to\infty \tag{C.1}$$

Die folgende Proposition zeigt nun, daß die Bedingung (5.6) auch für Speicherung von Zufallsmustern, die gemäß $Z(p)$ mit vorgegebener Aktivierung erzeugt wurden, das Hifi-Qualitätskriterium ausreichend ist.

Proposition 3
Es gilt für die binäre Hebbspeicherung von Inhaltsmuster, die mit Zufallsauswahlprozess Z(p) aus Abschnitt 3.4, also mit vorgegebener Aktivierung p erzeugt wurden:

Ist für Muster mit $l = np$ Einsen die Bedingung (5.6) erfüllt, so sind gleichzeitig für alle l die Bedingungen (C.1) erfüllt.

Beweis:
Einschrittretrieval mit binärer Hebbregel hat für Inhaltsmuster mit l Einsen als Adressmuster an einer Komponente des Ausgangsmusters die Fehlerwahrscheinlichkeiten:

$$\tilde{\alpha}_l = \sum_{l'=\Theta}^{l} B(l, p_0, l') \quad , \quad \tilde{\beta}_l = H\left[l - \Theta\right].$$

$H\left[x\right]$ ist die Heavisyde-Funktion und $B(n,p,l) := \binom{n}{l} p^l (1-p)^{n-l}$. Mit den Beziehungen (5.12) und (5.14) erhält man damit

$$p_{um}^l \approx 1 - (1 - \tilde{\alpha}_l)^{n-k} (1 - \tilde{\beta}_l)^k.$$

Wählen wir den Schwellenwert $\Theta = ck$ mit der Konstanten $c < 1$ nahe bei Eins und setzen (4.27) ein, so erhalten wir für (C.1)

$$\sum_{l=0}^{ck} B(n,p,l) + \sum_{l=ck}^{n} B(n,p,l) p_{um}^l =: A + B.$$

Der erste Term auf der linken Seite, hier mit A bezeichnet, kann mit der Gaußapproximation abgeschätzt werden: $A \approx G(\sqrt{k}(1-c))$, er verschwindet nach Prop. 1 aus Anhang B für $\ln[n]\,(1-c) \to \infty$. Da die Binomialverteilung symmetrisch um $l = k$ ist, liefern die Summanden ab $k(2-c)$ aufwärts im zweiten Term keinen Beitrag und können daher abgeschnitten werden: $B = \sum_{l=ck}^{k(2-c)} B(n,p,l) p_{um}^l$. Die Fehlerwahrscheinlichkeit p_{um}^l nimmt ihren maximalen Wert für $l = k(2-c)$ an. Mit Hilfe einer Chernoffapproximation für die Binomialsumme in $\tilde{\alpha}_l$ erhält man als Abschätzung des zweiten Terms: $B \leq p_{um}^{k(2-c)} \approx n\tilde{\alpha}_{k(2-c)} = n \exp[ck \ln[p_0]]$.

Fazit: Wenn c langsam genug gegen Eins geht, um noch das Verschwinden des Termes A zu gewährleisten, dann ist durch die Bedingung (5.6) das Verschwinden des Terms B sichergestellt, so daß (C.1) erfüllt ist. ♢

Wir beschränken uns hier auf die Darstellung des Beweises für die binäre Hebbregel, mutatis mutandis läßt sich Proposition 3 auch für inkrementelle Hebbregel zeigen.

Um unsere Ergebnisse mit den Berechnungen in den zitierten Arbeiten zu vergleichen, wurde ein fester Schwellenwert für alle Eingangsmuster gewählt. Die Retrievalqualität im endlichen Modell verbessert sich noch, wenn die Schwelle gleich der Anzahl der aktivierten Neurone für jedes individuelle Eingangsmuster gesetzt wird: $\Theta = \sharp\{i : x_i = 1\}$. Im thermodynamischen Limes verändert eine solche Variation der Schwellenadjustierung die asymptotischen Kapazitätswerte jedoch nicht.

Anhang D

Liste oft verwendeter Symbole

a	Off Aktivitätswert eines Musters
A	Ausgangskapazität
α	Fehlerwahrscheinlichkeit in Off-Neuronen eines binären Musters
B_x	Wertebereich einer diskreten Zufallsvariable
β	Fehlerwahrscheinlichkeit in On-Neuronen eines binären Musters
c	Synaptische Matrix bei inkrementeller Speicherung
$\bar{c}$	Synaptische Matrix bei binärer Speicherung
CA	Schwellenstrategie
d_{on}	Dendritisches Potential am On-Neuron
d_{off}	Dendritisches Potential am Off-Neuron
f	Verhältnis falscher zu richtiger a-Komponenten in einem binären Muster
F	Fixpunktkapazität
$\mathcal{F}$	Menge aller Fixpunkte
g	Verhältnis falscher zu richtiger 1-Komponenten in einem binären Muster
H	Hebbregel (Synaptische Änderungsvorschrift)
$\mathcal{H}(x)$	Energiefunktion des rückgekoppelten Speichers
$i(x)$	Shannon Informationsfunktion
$I(x \mid \tilde{x})$	Bedingte Information in $\tilde{x}$ über x
k	Anzahl von 1-Komponenten in den Inhaltsmuster
K	Korrekturkapazität
M^*	Maximale Anzahl von gespeicherten Inhaltsmuster, mit der das Qualitätskriterium noch erfüllt ist
$\mathcal{M}_k$	Menge aller Muster mit k 1-Komponenten
m_{on}	Anzahl der vorhandenen oder in einem Retrievalschritt hinzugefügten On-Neurone in einem Muster
m_{off}	Anzahl der fehlerhaften Off-Neurone in einem Muster
μ	Aktivität eines Musters
p	Wahrscheinlichkeit, mit der eine Musterkomponente eines Inhaltsmusters On-Neuron ist
P	