

E. von Goldammer und R. Kaehr [*]

'Lernen' in Maschinen und lebenden Systemen

neokonnektionismus – machine learning – neurocomputer

Einleitung

Ein Ziel der Gehirnforschung besteht sicherlich darin, anhand von neurophysiologischen Erkenntnissen Aufschlüsse über die 'Informationsverarbeitung' und 'Intelligenz' lebender Systeme zu erhalten, in der Hoffnung, eines Tages möglicherweise Aussagen über das Verhalten von Tier und/oder Mensch aus der Struktur und der Funktion neuronaler Netzwerke ableiten zu können.

Die Konstruktion von Robotern, die sich in einer 'unstrukturierten Umgebung' bewegen können, ist das erklärte Ziel der Robotik [1]. Dies setzt aber notwendigerweise 'lernfähige Maschinen' voraus. 'Lernfähigkeit' ist jedoch eines der wesentlichen Merkmale 'natürlicher Intelligenz' und damit stellt der gesamte Bereich des 'machine learning' einen integralen Bestandteil der Erforschung 'künstlicher Intelligenz (KI)' dar.

Trotz dieser recht unterschiedlichen Zielsetzungen existiert zwischen beiden Gebieten ein nicht zu übersehender Zusammenhang, der sich in einer wechselseitigen Beeinflussung beider Wissenschaftsbereiche manifestiert. Einen Höhepunkt in dieser Wechselbeziehung stellt zweifellos die Logifizierung neuronaler Aktivitäten durch "*A Logical Calculus of the Ideas Immanent in Nervous Activity*" von McCulloch & Pitts [2] aus dem Jahre 1943 dar. Ein Modell, das bis zum heutigen Tag als Paradigma in der Gehirnforschung fungiert. Darüber hinaus stellt diese Studie nicht nur den Ausgangspunkt für von Neumann's Arbeit [3] "*The General Logic Theory of Automata*" aus dem Jahre 1951 dar, die zur Grundlage der heute verbreitetsten sog. 'Von Neumann Maschinen' – den seriell arbeitenden Computern – wurde, sondern markiert gleichzeitig auch den Beginn der Forschungen auf dem Gebiet des sog. 'machine learning'. Hier sind u.a. auch von Neumann's Arbeiten über die "*Theorie selbstreproduzierender Automaten*" [4] zu nennen und die Beschreibung der Konzeption des 'Perceptrons' von Rosenblatt aus dem Jahre 1962 [5]. Eine Sammlung sämtlicher wichtiger Texte dieser Entwicklung findet sich in dem Buch von Anderson [6]. Heute hat das Gebiet des 'machine learning' in der Öffentlichkeit vor allem durch Schlagworte wie 'Neurocomputer', 'Neokonnektionismus', usw. eine gewisse Publizität erlangt.

Ein wesentliches Merkmal der Konzeption der 'Neurocomputer' ist deren 'Lernfähigkeit'. Es erhebt sich mithin die Frage, ob das hochgesteckte Ziel der Neuroinformatik bzw. der Robotik mit Hilfe dieses Lernkonzeptes realisierbar ist.

Lernen in neuromorphen Netzen

Die folgenden Abschnitte erheben nicht den Anspruch einen Überblick über das Prinzip des 'neuro-computing' zu geben, sondern haben zum Ziel, einen bestimmten Aspekt von 'Lernen' zu beleuchten und das Resultat in Relation zu Begriffen

* veröffentlicht in: DESIGN & ELEKTRONIK, Ausgabe 6 vom 21. März 1989, S. 146-151.

wie 'adaptives Lernen', 'adaptive Filter', 'Lernen ohne Lehrer', usw., wie sie im Zusammenhang mit neuromorphen Netzen benutzt werden, zu setzen.

Das Hauptmerkmal dieser Netze ergibt sich durch das Muster der Verbindungen ('Konnektionen' oder 'Synapsen') zwischen einzelnen 'Prozessorelementen' (PE). Die PE's bzw. ihre Software-Simulationen modellieren in den neuromorphen Netzen einige der wesentlichen Funktionen der Neuronen im Nervensystem. Die Hard- und/oder Software-Topologie der Netze zeichnet sich darüber hinaus durch eine hohe Parallelität der Datenströme aus. Durch die Implementierung von 'feedback'-Algorithmen, den sog. 'Lernregeln', sind neuromorphe Rechenetze besonders geeignet zur Erkennung durch Rauschen gestörter Muster. Ohne an dieser Stelle auf Details näher einzugehen, sei zur Topologie derartiger Netze kurz folgendes angeführt: Sie werden aus Schichten von PE's gebildet, die unterschiedliche Funktionen im Netz erfüllen. Man spricht daher von 'input'- bzw. 'output'-Schichten ('layer') und von sog. versteckten ('hidden') Schichten. Die Anzahl der PE's in den einzelnen Schichten hängt von der jeweiligen Konstruktion des Netzes ab. Das charakteristische funktionale Merkmal derartiger Rechenetze ist die variable Verbindungsstärke (Konnektionsstärke) zwischen den einzelnen PE's in den unterschiedlichen Schichten. Die bereits erwähnten Lernregeln gehen dabei einmal auf Arbeiten von Hebb [7] aus dem Jahre 1949 und Widrow & Hoff (1960) [8] zurück. Diese Regeln, die sich durch relativ einfache mathematische Beziehungen darstellen lassen, besagen in Worten folgendes

a) HEBB'sche REGEL

Werden zwei verbundene PE's gleichzeitig angeregt, dann nimmt die Gewichtsfunktion ihrer Verbindung (Konnektion) zu.

Dies wird als 'nicht unterwiesenes' oder 'Lernen ohne Lehrer' bezeichnet.

b) WIDROW-HOFF oder DELTA-REGEL

Das Netz akzeptiert Eingangsmuster und führt Transformationen durch. Die Ausgangssignale werden mit einer 'Schablone' verglichen und die Konnektionen der PE's nach bestimmten Algorithmen auf optimale Übereinstimmung zwischen 'output' und 'Schablone' einjustiert.

Diese Prozedur benötigt ein 'feedback' in Form der Schablone oder irgendeines 'Lehrers'.

Die Hebb'sche Regel wird auch als 'Selbstorganisations-Regel' bezeichnet: auf diesen Begriff wird später zurückzukommen sein. Zur 'Hebb'schen Lernkonzeption' gehören auch die 'Hopfield-', 'Boltzmann-' oder 'Cauchy'-Modelle [6], die alle mehr oder weniger physikalischen Ursprungs sind und ihr Pendant jeweils in der Physik (Spingläser, etc.) besitzen, im Gegensatz zu den 'NON-Hebbschen' Lernregeln [9], die stärker biologisch fundiert sind und eine kooperative, kontextabhängige Konzeption darstellen.

Obwohl die heuristischen Möglichkeiten infolge des implementierten 'feedback' bei den neuromorphen Netzen ausgezeichnet sind und sich diese Netze daher hervorragend für die Erkennung unvollständiger Muster eignen, werden die Hoffnungen, mit Hilfe derartiger Rechenetze eines Tages beispielsweise Roboter konstruieren zu können, die sich in einer 'unstrukturierten' Umgebung bewegen kön-

nen, nicht in Erfüllung gehen – jedenfalls nicht auf der Basis der bisher verwendeten Algorithmen. Zum besseren Verständnis sei hier angemerkt, daß eine 'strukturierte Umgebung' für einen Roboter, wie sie in der Robotik heute üblich ist, im kybernetischen Sinne keine 'Umgebung' darstellt, d.h. hier ist die 'Umgebung' ein Teil des Robots, da sie sozusagen einprogrammiert ist. Roboter, die mit entsprechenden Sensorsystemen und dem Apparat neuromorpher Netze ausgestattet sind, werden also prinzipiell nicht in der Lage sein, sich in einer unstrukturierten, d.h. einer mehr oder weniger beliebigen Umgebung, frei bewegen zu können. Ein System, das diese Fähigkeit besitzt, und dazu gehören ganz offensichtlich lebende Systeme, ist prinzipiell in der Lage, zwischen sich (selbst) und der Umgebung unterscheiden zu können (Selbstreferentialität), denn das bedeutet ja gerade, "sich frei in einer beliebigen (= unstrukturierten) Umgebung zu bewegen". Für einen derartigen Vorgang, den man möglicherweise als einen 'intelligenten Prozeß' zu bezeichnen geneigt ist, ist ein 'Lernen' notwendig (im Sinne einer notwendigen, jedoch nicht unbedingt hinreichenden Bedingung), das von dem Anthropologen und Kybernetiker Bateson als 'Lernen_II' oder 'Lernen lernen' bezeichnet wurde, während es sich bei den übrigen bisher beschriebenen Lernprozessen in der Terminologie Bateson's um 'Lernen_I' (Hebb) bzw. 'Lernen_0' (Widrow/Hoff) handelt.

Da diese Unterscheidung von 'Lernen' in der (Fach)Literatur über neuromorphe Netze im allgemeinen nicht gemacht und daher auch nicht diskutiert wird, sollen im folgenden zunächst einige der Gedankengänge aus den Arbeiten von Bateson [10] nachgezeichnet und im Anschluß daran die Konsequenzen für das Konzept des 'machine learning' diskutiert werden. Es sei hier nur angemerkt, daß es sich dabei um eine kybernetische Lerntheorie handelt; auf die Ergebnisse der Lernpsychologie soll hier nicht eingegangen werden.

Die logischen Kategorien von Lernen und Kommunikation

So lautet der Titel eines Aufsatzes von Bateson [10] aus dem Jahre 1964, in dem eine logische Analyse von Kommunikationsprozessen, zu denen das 'Lernen' gehört, mit Hilfe der logischen Typenlehre von Russell vorgenommen wird.

Es erscheint angebracht, an dieser Stelle kurz zu rekapitulieren, worum es bei der Theorie der logischen Typen geht. Sie besagt, daß keine Menge in der formalen (klassischen) Logik oder im mathematischen Diskurs Element ihrer selbst sein kann, daß eine Menge von Mengen nicht eine der Mengen sein kann, die ihre Elemente sind. Hauptzweck der Russell'schen Typenlehre ist es, imprädikative, d.h. zirkuläre Definitionen durch das sog. 'vicious circle principle' (Zirkelprinzip) auszuschließen: Ein Begriff, der unter Bezugnahme auf eine Gesamtheit definiert ist, darf nicht selbst dieser Gesamtheit angehören.

Die Typentheorie unterscheidet nicht nur zwischen Element und Gesamtheit (Menge) bzw. zwischen Individuum und Prädikat, sondern auch zwischen verschiedenen Typen. und zwar so, daß zwischen Element und Menge (Individuum und Prädikat) immer ein Typenunterschied eingehalten werden muß. Dadurch ist eine selbstrückbezügliche Begriffsbildung nicht möglich. Die Einhaltung des Zirkelprinzips garantiert eine antinomienfreie Formalisierung. Da die Typenbildung unabgeschlossen ist, wird das Problem der Selbstrückbezüglichkeit, das für jeden kognitiven Prozeß konstitutiv ist, ins 'Unendliche' verschoben. Wenn also die

'Umgebung' nicht Teil des Robots, d.h. nicht einprogrammiert sein soll, so muß dieser zwischen sich und seiner Umgebung eine Unterscheidung treffen können – dies ist jedoch ein selbstrückbezüglicher Prozeß, worauf von seiten der Philosophie seit Kant hingewiesen wird.

Auf den ersten Blick erscheint dies alles zwar recht trivial und offensichtlich, dennoch wird es gerade bei der Beschreibung von Verhaltensweisen lebender Systeme sehr häufig übersehen, so daß Irrtümer der logischen Typisierung auftreten. Da also eine Menge der Stühle kein Stuhl ist, folgt entsprechend, daß auch die Menge der Nichtstühle kein Nichtstuhl ist. Schließlich besagt die Theorie, daß bei einem Verstoß gegen diese einfachen Regeln des formalen Diskurses Paradoxien erzeugt werden, die für das formale System destruktiv sind.

Schon an dieser Stelle erkennt man, daß das Problem (Simulation von 'Lernen' in Maschinen) möglicherweise bereits darin besteht, das 'Problem' erst einmal logisch adäquat zu formulieren – das ist *das Problem*.

Folgen wir zunächst den Argumenten von Bateson, dann wird in gewöhnlicher nichttechnischer Redeweise das Wort 'Lernen' häufig auf Phänomene angewandt, welche von ihm als 'Lernen_null' bezeichnet werden, und die man am einfachsten in Analogie zu einer elektronischen Schaltung beschreiben kann. Hier ist die Schaltstruktur, insbesondere der Algorithmus selbst, keiner Veränderung unterworfen und die kausalen Verknüpfungen zwischen 'Reiz' und 'Reaktion' ('input' und 'output') sind sozusagen "eingelötet". Ein Beispiel dafür stellt die eingangszitierte Widrow/Hoff'sche Lernregel dar. Auf das von Bateson beschriebene intellektuell recht interessante Beispiel für 'Lernen_null', den hypothetischen 'von Neumann'schen Spieler', für den 'Versuch und Irrtum' nicht existieren (!), sei hier nicht näher eingegangen. Es ist jedoch von entscheidender Bedeutung für die Begriffsklärung, daß auf der Ebene des 'Lernen_null' definitionsgemäß der 'Irrtum' ausgeschlossen ist und damit das Prinzip von 'Versuch und Irrtum' nicht zum Repertoire dieses Lernprozesses gehört.

Wenn man nun die Vorstellung akzeptiert, daß alles Lernen (außer dem 'Lernen_null') in einem gewissen Maße stochastisch abläuft, d.h. Komponenten von 'Versuch und Irrtum' enthält, dann folgt, daß sich eine mögliche Ordnung der Lernprozesse auf eine hierarchische Klassifizierung der Irrtumstypen, die in den vielfältigen Lernprozessen jeweils korrigiert werden sollen, stützen kann. D.h., Bateson benützt die deskriptive Typentheorie zum Zweck einer wissenschaftlichen Präzisierung in Verbindung mit den Russell'schen Typen – eine Beschreibung, die es ihm ermöglicht, Fehlverhalten in Lernprozessen als Typenverwechslung bzw. als Verstoß gegen das Zirkelprinzip zu interpretieren.

'Lernen_null' wird dann zur Bezeichnung für all jene Prozesse, die nicht der Berichtigung durch Versuch und Irrtum unterworfen sind. 'Lernen_I' gebraucht er als Etikett für die Aspekte des Lernens, bei denen die Zurücknahme der Wahl innerhalb einer unveränderlichen Menge von Alternativen charakteristisch ist. 'Lernen_II' steht als Etikett für alle jene Lernprozesse, die sich durch eine Revision der Menge, aus der die Wahl getroffen werden soll, auszeichnen; und so weiter.

Als ein Beispiel für 'Lernen_I' führt Bateson die Pawlow'sche Konditionierung an, wobei der Begriff des 'wiederholbaren Kontextes' für die Begriffsbildung eine notwendige Voraussetzung darstellt. D.h. die Wiederholbarkeit eines (Lern)Kontextes ist in seiner Beschreibung von Lernvorgängen, die 'Lernen als Veränderung'

auffaßt, eine zwingend notwendige Bedingung. Ohne diese Annahme eines wiederholbaren Kontextes wäre alles 'Lernen' von einem Typ, nämlich 'Lernen_null', so daß alle Fragen der Art *„Ist dieses Verhalten 'erlernt' oder 'angeboren'?* zugunsten der Genetik beantwortet werden müßten – was unseren Erfahrungen sicherlich widerspricht.

Als weiteres Beispiel für 'Lernen_I' sei aus dem technischen Bereich das Lernen gemäß der Hebb'schen Regel angeführt, das auch häufig mit dem Begriff der 'Selbstorganisation' in Verbindung gebracht wird. In der Terminologie Bateson's bedeutet dies, daß auch der Begriff der 'Selbstorganisation' in diesem Kontext vom gleichen logischen Typus sein muß, wie das 'Lernen_I'. Es sei an dieser Stelle darauf hingewiesen, daß gerade bei der Verwendung des heute recht in Mode gekommenen Begriffs der 'Selbstorganisation' die eingangs bereits erwähnten Irrtümer der logischen Typisierung im Diskurs sehr häufig gemacht werden.

Während also 'Lernen_I' die Veränderung durch Korrektur von Irrtümern über die Auswahl innerhalb einer Menge von Alternativen ist, stellt 'Lernen_II' die Veränderung im Prozeß des 'Lernen_II' dar. Hier tritt also, um mit Bateson zu sprechen, eine korrigierende Veränderung in der Menge von Alternativen, unter denen die Auswahl getroffen wird, auf, was etwas technischer gesprochen der Veränderung des Algorithmus (durch das System selbst) entspricht. D.h. die Veränderung betrifft nicht nur die Daten des Algorithmus, die Operanden (das wäre 'Lernen_I'), sondern es handelt sich beim 'Lernen_II' und allen höheren Lerntypen um eine Variation der Operatoren, also der Lernregeln durch das System (!). Die Veränderung der Operatoren beim 'Lernen_II' entspricht einer Zirkularität ('Lernen des Lernen') und verstößt gegen das Zirkelprinzip der Typentheorie, während die Veränderung der Operanden beim 'Lernen_I' im formalen Sinn einer Rekursion gleichkommt.

'Lernen_III' ist in dieser Terminologie dann folgerichtig die Veränderung im Prozeß des 'Lernen_II', z.B. eine korrigierende Veränderung im System der Mengen von Alternativen, unter denen die Auswahl getroffen wird. Hier kann nicht weiter auf Prozesse von 'Lernen_III' (oder gar 'Lernen_IV') eingegangen werden. Es sei nur angemerkt, daß entsprechend der Typentheorie die Lerntheorie Bateson's nach oben nicht abgeschlossen ist.

'Lernen_II' läßt sich u.a. experimentell durch das sog. 'Umkehrungs-Lernen' feststellen. Bei diesen Experimenten lernt das System zunächst immer eine binäre Entscheidung ('Lernen_I'). Wenn diese als Kriterium gelernt ist, wird die Bedeutung der Reize umgekehrt. D.h. wenn anfänglich $X \rightarrow R_1$ und $Y \rightarrow R_2$ bedeutete, dann bedeutet nach Umkehrung $X \rightarrow R_2$ und $Y \rightarrow R_1$. Wenn dies als Kriterium gelernt ist, dann werden die Bedeutungen wiederum umgekehrt. Die entscheidende Frage bei diesem Experiment ist: Lernt das Subjekt etwas über die Umkehrung? D.h. erreicht das System das Kriterium nach einer Reihe von Umkehrungen mit weniger Versuchen?

Die derart gestellte Frage ist eindeutig von einem höheren logischen Typ als die Fragen zum 'Lernen_I'. Der Inhalt dieser Frage ist völlig analog zu Russell's Relation zwischen 'Menge' und 'Menge von Mengen'.

Ein technisches Beispiel für 'Lernen_II' kann hier nicht gegeben werden, da eine Maschine, die ihren eigenen Algorithmus (ohne 'Lehrer') erstellt, bisher nicht kon-

struiert wurde. Alles bis heute beschriebene maschinelle Lernen ist entweder vom Typ 'null' oder 'eins'. Da jedoch Lernprozesse phänomenologisch gesehen in der oben beschriebenen Art klassifiziert werden müssen, stößt man hier auf ein grundlegendes logisches Problem, nämlich der Modellierung von Übergängen zwischen unterschiedlichen logischen Typen, um in der Terminologie Bateson's zu bleiben. Auf dieses Problem weist Bateson in seinen Aufsätzen bereits deutlich hin und weiter unten wird darauf zurückzukommen sein. Was Bateson nicht diskutiert, ist die Problematik, die sich in seinem Modell der Beschreibung von Lernprozessen durch das sog. Reduktions-Theorem (vgl.: Quine [11]) ergibt, wonach sich ein System von n Typen auf nur einen Typ reduzieren läßt. so daß sich allein aufgrund dieses Theorems (in der Terminologie Bateson's) logisch formal gesehen alles 'Lernen' auf einen Typ – also 'Lernen_null' – reduzieren läßt, was seine gesamte Analyse natürlich *ad absurdum* führen würde.

Fassen wir an dieser Stelle kurz zusammen:

1. Es ist sicherlich unbestritten, daß Lernprozesse beobachtet und zu ihrer Beschreibung in Kategorien eingeteilt werden können, die von unterschiedlichem Typ sind.
2. Es existieren innerhalb der klassischen Logik keine Operatoren, die eine Vermittlung zwischen unterschiedlichen logischen Typen zulassen.
3. Formal läßt sich ein System aus n Typen immer auf einen Typ reduzieren (Typenreduktion), wodurch die einzelnen Prozesse logisch strukturgleich, d.h. homogenisiert werden.

Hier besteht also ein offensichtlicher Widerspruch zwischen dem phänomenologischen Bestand der Lerntheorie (Selbstrückbezüglichkeit, Irreduzibilität der Lern-typen) und der applizierten Methode (Typentheorie) ihrer Szientifizierung, d.h. das Problem liegt in der Verknüpfung der Lerntheorie mit der Typentheorie und nicht am phänomenologischen Bestand der Lerntheorie.

Definition einer 'POLYLOGISCHEN MASCHINE'

Da das Gebiet des 'machine learning' ganz offensichtlich als integraler Bestandteil der KI-Forschung angesehen werden muß, sind logische Analysen von Lernprozessen von fundamentaler wissenschaftlicher Bedeutung für diesen Bereich der Ingenieurwissenschaften. Es ist daher um so erstaunlicher, daß diese Thematik weitgehend ignoriert und wenn überhaupt, dann allenfalls von philosophischer Seite (z.B.: Dreyfus [12]) aufgegriffen wird, und dies, obwohl von amerikanischen Kybernetikern grundlegende Arbeiten über die Theorie der Kognition angefertigt wurden. Beschäftigt man sich mit diesen Arbeiten und das gilt auch für die bereits erwähnten Aufsätze von Bateson, dann fällt auf, daß ein Aspekt immer wieder auftaucht, nämlich der explizite (von Foerster [13], McCulloch [14], Hofstadter [15], Varela [16], u.a.) oder implizite (Bateson [10], Maturana [17]) Hinweis auf die Grenzen der klassischen Logik und ihrer Typentheorie als Hilfsmittel zur adäquaten Beschreibung von Problemen aus dem Bereich der Kognitionswissenschaften.

Wie muß eine Maschine konzipiert sein, in welcher Prozesse wie der des 'Lernen_II' modelliert werden können? Eine derartige Maschine soll im folgenden als 'poly-logische Maschine (PLM)' definiert werden:

Eine PLM soll in der Lage sein, logisch ablaufende Prozesse (logische Operationen) auszuführen und parallel dazu jeden einzelnen Schritt eines derartigen Prozesses zu analysieren und die Resultate der Analyse in Wechselbeziehung zu den Schritten der Prozesse zu setzen, um diese gegebenenfalls steuernd zu korrigieren, d.h. zu verändern.

Dies entspricht der Konzeption einer Maschine, bei der die Operatoren des einen Prozesses simultan als Operanden von Operatoren eines anderen Prozesses fungieren können. D.h. es besteht hier folgende Simultaneität von Unterscheidungen (Differenzen):

$$\left. \begin{array}{l} \text{Unterscheidung}_1 : (\text{Operator } \underline{O} \mid \text{Operand } O) \\ \text{Unterscheidung}_2 : (\text{Operator } \underline{O} \mid (\text{Operator } \underline{O} \mid \text{Operand } O)) \end{array} \right\} \quad (1)$$

Gemäß der Russell'schen Typentheorie muß zwischen Operator und Operand ein Typenunterschied eingehalten werden, um einer Zirkularität zu entgehen. Das bedeutet aber, daß derartige, simultan, ablaufende Prozesse mit Hilfe der logischen Typen nicht modellierbar sind, da es keine Möglichkeiten für die Darstellung von Übergängen zwischen einzelnen Typen gibt außer dem der Über- und Unterordnung (Hierarchie). D.h. es fehlt die Vermittlung zwischen den Typen.

Nun kann man natürlich fragen, ob es überhaupt sinnvoll ist, eine Maschine dieser Art zu konzipieren, wenn dabei derartige Schwierigkeiten auftauchen.

Eingangs wurde bereits darauf hingewiesen, daß sich lebende Systeme durch ihre kognitiven Fähigkeiten auszeichnen, d.h. beispielsweise zwischen sich und ihrer Umgebung unterscheiden können. Genau diese Prozesse sind strukturgleich zu den Prozessen in der hier definierten 'poly-logischen Maschine, d.h. auch hier treten simultan folgende Differenzen auf:

$$\left. \begin{array}{l} \text{Unterscheidung}_1 : (\text{System } \underline{O} \mid \text{Umgebung } O) \\ \text{Unterscheidung}_2 : (\text{System } \underline{O} \mid (\text{System } \underline{O} \mid \text{Umgebung } O)) \end{array} \right\} \quad (2)$$

Der **kognitive Prozeß** konstituiert sich erst durch den selbstrückbezüglichen oder selbstreferentiellen Prozeß der Reflexion, bei dem im System ein Abbild von System und Umgebung erzeugt wird.

Kehren wir an dieser Stelle nochmals zum Prozeß des 'Lernens' zurück und betrachten die Situation des sog. 'Umkehrungs-Lernens' als ein Beispiel für den Prozeß von 'Lernen_II' oder 'Lernen lernen', um wieder die Terminologie von Bateson zu benutzen. Bei diesem Prozeß wurde einmal die Unterscheidung $X \rightarrow R_1$ und

$Y \rightarrow R_2$ gelernt (mit X, R für Reiz und Reaktion des Systems). Wir benutzen dafür folgende Schreibweise:

$$\underline{O}_1(X \rightarrow R_1, Y \rightarrow R_2) = \underline{O}_1(O_1) = \text{Operator}(\text{Operand}) \tag{3a}$$

und für die umgekehrte Lernsituation entsprechend

$$\underline{O}_2(X \rightarrow R_2, Y \rightarrow R_1) = \underline{O}_2(O_2) \tag{3b}$$

Die entscheidende Frage bei diesen Experimenten ist: lernt das System etwas über die Umkehrung? Damit stehen wir jedoch wiederum vor einer Situation, die der aus (1) bzw. (2) völlig strukturgleich ist, d.h.:

$$\left\{ \begin{array}{l} \text{Unterscheidung}_1 : (\underline{O}_1 | O_1) \ \& \ (\underline{O}_2 | O_2) \quad : \text{Lernen_I} \\ \text{Unterscheidung}_2 : (\underline{O} | (\underline{O}_1 | O_1) \ \& \ (\underline{O}_2 | O_2)) \quad : \text{Lernen_II} \end{array} \right\} \tag{4}$$

Wenn also derartig strukturierte Prozesse in einer Maschine simuliert werden sollen und das ist letztendlich das Bestreben eines ingenieurwissenschaftlichen Programms wie dem des 'machine learning', dann muß offensichtlich im Sinne der Konzeption einer 'poly-logischen Maschine' Zirkularität parallel distribuiert über das Verhältnis von Operator/Operand sein und umgekehrt. Damit ist eine Parallelität gefordert, die sich erst durch die Zirkularität von Operator und Operand konstituiert. Dies soll im folgenden an einer Skizze (Abb.1) veranschaulicht werden.

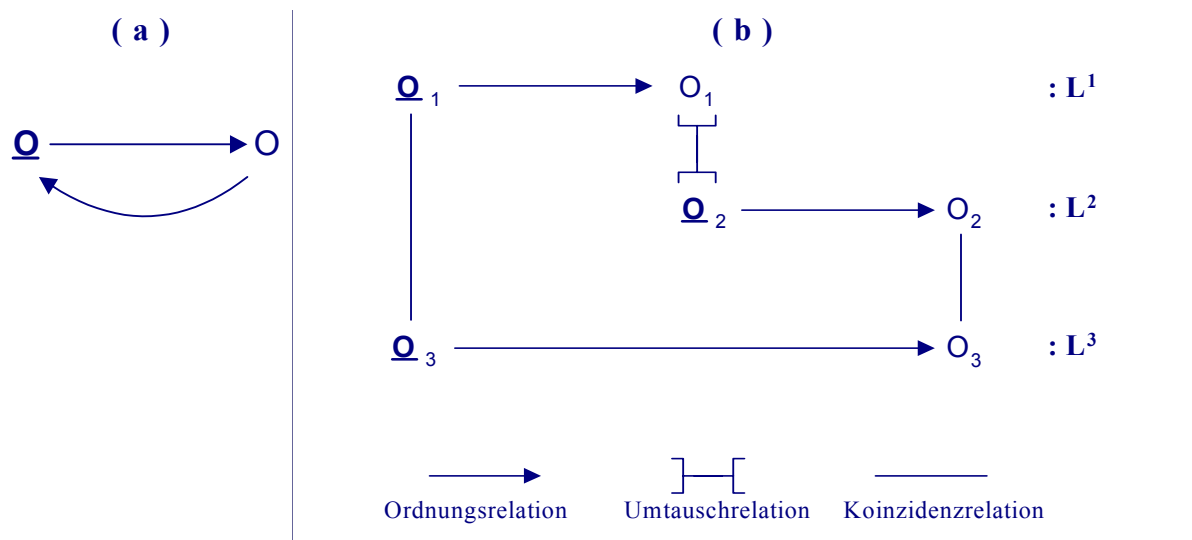


Abb. 1 : Drei Logik-Systeme $L^{(3)} = (L_1, L_2, L_3)$
 (a) Zirkularität durch Selbstrückbezüglichkeit in einem klassischen mono-kontexturalen System;
 (b) distribuierte Zirkularität und Parallelität in einem 'poly-logischen'System.

Im Rahmen der Typentheorie sind zwischen den einzelnen Typen, wie bereits erwähnt, nur Ordnungsrelationen im Sinne der Über- bzw. Unterordnung (Hierarchie) vorhanden. Damit ist jedoch eine Modellierung der 'Selbstrückbezüglichkeit' prinzipiell ausgeschlossen, denn diese würde einen eindeutigen Verstoß gegen das Zirkularitätsprinzip darstellen. In der Abb._1a ist die zirkuläre Situation (Selbstrückbezüglichkeit) graphisch nochmals dargestellt. Auf der anderen Seite sind in

der Abb._1b drei logische Bereiche $L^{1, 2, 3}$ derart dargestellt, daß sich hier Zirkularität über diese Bereiche verteilt, wenn man die Begriffe beim Übergang von einem Bereich (Domäne) zum anderen beibehält. Hier wird eine neue und höhere Form von 'Zirkularität' eingeführt (Chiasmus). Das Verhältnis von Operator und Operand distribuiert sich auf zwei logische Domänen und entgeht der Zirkularität, wenn die Prozesse jeweils unterschieden werden. Man bezeichnet dieses Wechselspiel zwischen verschiedenen Bereichen (bei denen die Zirkularität von \underline{O} und \bar{O} verteilt ist) auch als 'Heterarchie' (Nebenordnung; *heteros*: das andere). Da es sich hier nicht mehr um die Russell'sche Typentheorie handelt, sprechen wir im folgenden nicht mehr von 'Typen', sondern von 'Kontexturen'. Eine Kontextur ist jeweils eine logische Domäne, in der alle klassischen logischen Regeln ihre volle Gültigkeit besitzen, d.h. eine Kontextur ist ein Bereich, der durch die 2-wertige klassische Logik oder eine klassisch mehrwertige Logik beschrieben werden kann.

Polykontexturalität entsteht dadurch, daß zwischen je zwei Werten eine klassische Logik definiert ist. Bei drei Werten, wobei der dritte Wert nicht zwischen 'wahr' und 'falsch' sondern außerhalb von 'wahr' und 'falsch' steht, erhält man drei 2-wertige Logiken, denen 3 Kontexturen zugeordnet sind. Bei vier Werten erhält man sechs Logiken bei m Werten " $2_{\text{über}}_m$ " Logiken. Wesentlich dabei ist nun, daß die so definierten Logiken oder Kontexturen nicht isoliert nebeneinander stehen, sondern durch Ordnungs- und Umtauschrelationen miteinander vermittelt sind und in dieser Vermittlung liegt der fundamentale Unterschied zur Russell'schen Typentheorie. Genau dieser Mechanismus von Simultaneität und Vermittlung verhindert eine Reduktion im Sinne der Typentheorie.

Während sich Heterarchie in diesem Modell inter-kontextural konstituiert, ist innerhalb einer logischen Domäne, d.h. intra-kontextural jeder Prozeß hierarchisch strukturiert. Daraus wird ersichtlich, daß der Begriff der 'Heterarchie' den der 'Hierarchie' nicht ausschließt, sondern distribuiert, denn die nebengeordneten, vermittelten Prozesse (oder Systeme) sind jeweils in sich hierarchisch strukturiert.

Es würde den Rahmen dieser Arbeit sprengen, die 'Theorie der Polykontexturallogik (PKL)', die von Günther [18] eingeführt und von Kaehr [19] in der Folgezeit weiterentwickelt wurde, an dieser Stelle ausführlicher darzustellen. Hier sollten lediglich die Möglichkeiten und die Grenzen bei der Simulation kognitiver Prozesse mit Hilfe neuromorpher Netze etwas ausgeleuchtet und in diesem Zusammenhang auf ein grundsätzliches wissenschaftslogisches Problem hingewiesen werden.

Da die Typentheorie infolge ihrer ausschließlich hierarchischen Struktur keine Vermittlung zwischen verschiedenen gleichrangigen Typen kennt, ist eine Modellierung von Simultaneität von Prozessen im oben diskutierten Sinne prinzipiell ausgeschlossen. Die Möglichkeit einer derartigen Modellierung ist jedoch als Grundvoraussetzung für jegliche parallele Architektur in Soft- und Hardware anzusehen. Die PKL stellt eine Basis für die Simulation derartiger Prozesse dar.

Infolge der Umtauschrelation zwischen den verschiedenen Kontexturen kennt die PKL keine Entsprechung zur 'Typenreduktion'; ein Reduktionsverfahren, durch das ein System von n Typen – genau wie die Mehrsortenlogik auf eine Einsortenlogik – ohne Verlust auf einen Typ reduziert werden kann. Selbstverständlich hat die Mehrsortenlogik bzw. die Typentheorie ihre praktischen Vorteile, leistet jedoch formal nichts, was nicht auch durch die Einsortenlogik geleistet wird.

Heterarchie und Hierarchie

Man hat in der Fachliteratur kaum zur Kenntnis genommen, daß kurz nach Erscheinen von "*A Logical Calculus ...*" [2] von McCulloch eine Studie mit dem Titel "*A Heterarchy of Values Determined by the Topology of Nervous Nets*" [20] erschienen ist, in der er auf die Notwendigkeit der Einführung NON-transitiver und heterarchischer Strukturen zur adäquaten Beschreibung neuronaler Aktivitäten hinweist. Heterarchische Strukturen entziehen sich jedoch wegen ihrer irreduziblen topologischen Distribution, wie dies oben bereits darzustellen versucht wurde, einer Logifizierung im klassisch-relationslogischen Rahmen. Auf diesen Sachverhalt hat wiederum McCulloch in seiner Arbeit "*What's in the Brain That Ink May Character?*" [14] deutlich hingewiesen.

Topologische Beschreibungen bieten einen Rahmen zur Beschreibung neuronaler Aktivitäten [21]. Ein topologisches Beschreibungskonzept, nämlich das der Ultrametritizität wurde zur Beschreibung physikalischer Systeme und in der Folgezeit auch für assoziative Speicher [22] diskutiert. Eine der Bedingungen für einen metrischen Raum ist durch die Dreiecksungleichung gegeben:

$$d(A,B)+d(B,C)\geq d(A,C) \quad (5)$$

Ultrametritizität ist durch die strengere Beziehung definiert:

$$d(A,C)\leq \max\{d(A,B),d(A,C)\} \quad (6)$$

Ultrametritizität korrespondiert unmittelbar mit hierarchischen Strukturen [23]; logisch betrachtet stellt die Beziehung (5) nichts anderes als die Transitivitätsrelation dar. Die bisher in der Literatur diskutierten neuromorphen Strukturen sind alle ultrametrisch und damit hierarchisch strukturiert, mit Ausnahme der "topologisch" beschriebenen Netze von McCulloch [20]. Eine logische Interpretation dieser McCulloch'schen Arbeit findet sich in Ref.[24].

Das Hebb'sche Modell der Selbstorganisation bezieht sich auf die Struktur der Daten (Operanden wie z.B.: elektrische Spannungen, Ströme, etc.) innerhalb des Netzes. Die interne Struktur der Datenorganisation ist selbstorganisierend. Das Verhältnis vom neuromorphen Netz (System) zu seiner Umgebung ist unabhängig von diesem so definierten Selbstorganisationsprinzip der Daten innerhalb des Netzes. D.h. diese neuromorphen Netze sind im kybernetischen Sinne Input/Output-Systeme, die Eigenschaft der Selbstorganisation besitzen. Solche Systeme sind ultrametrisch und streng hierarchisch organisiert – es gilt die klassisch-logische Transitivität. Für diese klassisch-kybernetischen Systeme sind daher die klassischen mathematischen Methoden und deren Logik zur Beschreibung ausreichend.

Heterarchisch strukturierte Systeme sind so charakterisiert, daß das Verhältnis eines solchen Netzes zu seiner Umgebung selbstorganisierend ist, d.h. hier verändert sich der Algorithmus des Systems bzw. die Operatoren, die das System definieren, und nicht nur die Daten. Derartige Netze gehören nicht zur Familie der Input/Output-Systeme, sondern zur Klasse der autonomen, in sich geschlossenen, heterarchischen Systeme, wie sie in der 'Theorie lebender Systeme' von Maturana & Varela [16, 17] als 'autopoietische Systeme' bezeichnet wurden. Solche Strukturen verlangen zur adäquaten formalen Darstellung eine neue Logik, wie sie durch die PKL vorgeschlagen wird.

Zusammenfassung

Die neuromorphen Netze klassischer Konzeption, die ihren Ausgangspunkt in der Arbeit von McCulloch & Pitts [2] haben, sind wegen ihrer ultrametrischen topologischen Struktur der Klasse der hierarchischen (mono-kontexturalen) Systeme zuzuordnen.

Die im vorliegenden Aufsatz besprochenen Systeme zeichnen sich durch simultane und vermittelte Parallelität ihrer Prozesse aus und stellen heterarchische Netze dar, in denen Selbstrückbezüglichkeiten antinomienfrei konzipiert werden können, die durch die PKL beschreibbar sind. Es war wiederum McCulloch [20], der als erster auf derartige Strukturen neuronaler Aktivitäten hingewiesen und somit den Begriff der 'Heterarchie' in die Computerwissenschaft eingeführt hat.

Systeme, in denen 'Differenzen von Differenzen' nicht-reduzierbarer Art (vgl. Gl. 1, 2, 4) vollzogen werden, sind nicht mehr durch das Identitätsprinzip charakterisierbar. Ihre formale Beschreibung stellt daher ein genuin logisch-strukturelles, und kein Problem der nicht-linearen Beschreibungsmethoden, wie beispielsweise im Bereich der 'Synergetik oder Chaostheorie' dar.

Literaturverzeichnis

- 1 / M.Brady, L.A.Gerhardt & H.F.Davidson (eds.),
"Robotics and Artificial Intelligence", Springer-Verlag, Heidelberg (1984).
- 2 / W.S.McCulloch & W.Pitts
A Logical Calculus of the Ideas Immanent in Nervous Activity, Bull.Math.Phys.
5 (1943) p.115-133; siehe auch Ref.14.
- 3 / J. von Neumann
The General and Logic Theory of Automata, in: Cerebral Mechanisms in
Behavior, The Hixon Symposium (L.A.Jeffress, ed.) John Wiley (1951), p.1-32.
- 4 / J. von Neumann
Theory of Self-Reproducing Automata, (A.Burks, ed.) Univ.III.Press (1966).
- 5 / F.Rosenblatt
The Perceptron: A Probabilistic Model for Information Storage and Organization
in Brain, Physiol.Rev. 65 (1958) p.386-408.
- 6 / J.A.Anderson & E.Rosenfeld
Neurocomputing, The MIT Press, Cambridge, MA (1988).
- 7 / D.O.Hebb
The Organization of Behavior, John Wiley (1949).
- 8 / B.Widrow & M.E.Hoff
Adaptive Switching Circuit, in: 1960 IRE WESCON Convention Record, Part 4,
p.96-104.
- 9 / F.Varela
Non-Hebbian Synaptic Rules, in: Cognitiva'85, CESTA, Paris (1985), p.751-753.
- 10/ G.Bateson
Ökologie des Geistes, Suhrkamp-Verlag (1983).
- 11/ W.V.O.Quine
On the Theory of Types, in: Journ.of Symbolic Logic 3 (1938) p.125-139.
- 12/ H.Dreyfus
What Computers Can't Do: A Critique of Artificial Intelligence, Harper + Row,

- N.Y. (1972).
- 13/ H. von Foerster
Sicht und Einsicht, Vieweg-Verlag (1985).
- 14/ W.S.McCulloch
Embodiments of Mind, The MIT Press, Cambridge, MA (1988).
- 15/ D.R.Hofstadter
Gödel, Escher, Bach: Ein endlos geflochtenes Band, Klett-Cotta (1985).
- 16/ F.Varela
Principles of Biological Autonomy, Elsevier North-Holland Publ. (1979).
- 17/ H.Maturana
Erkennen: Die Organisation und Verkörperung von Wirklichkeit, Vieweg-Verlag (1985).
- 18/ G.Günther
Beiträge zur Grundlegung einer operationsfähigen Dialekti, Vol.I-III, Felix
Meiner-Verlag, Hamburg (1980).
- 19/ R.Kaehr
in: Idee und Grundriß einer nicht-Aristotelischen Logik (G.Günther, ed.), Felix
Meiner-Verlag, Hamburg (1978).
- 20/ W.S.McCulloch
A Heterarchy of Values Determined by the Topology of Nervous Nets,
Bull.Math.Biophys. 7 (1945) p.89-93; siehe auch Ref. 14.
- 21/ P.Erdi
On the Ultrametric Structure of Semantic Memory: Scope and Limits, in:
Cybernetics and Systems (R. Trappel, ed.), Kluver Acad.Publ. (1988), p.329-336.
- 22/ T.Kohonen
An Introduction to Neural Computing. Neural Networks (INNS) 1(1988) p.3-16.
- 23/ R.Rammel, G.Toulouse & M.A.Virasoro
Ultrametricity for physicists, Rev.Mod.Phys. 58 (1986) p.765-788.
- 24/ R. Kaehr & E. von Goldammer
Again Computers and the Brain, Journ.Molecular Electronics 4 (1988) S31-S37.

The text was originally edited and rendered into PDF file for the e-journal <www.vordenker.de> by E. von Goldammer

Copyright 2005 vordenker.de

This material may be freely copied and reused, provided the author and sources are cited
a printable version may be obtained from webmaster@vordenker.de

vordenker

ISSN 1619-9324