

# Dynamische Systeme in der Kognitionswissenschaft

Herbert Jaeger

*GMD, St. Augustin*  
*email: herbert.jaeger@gmd.de*

*August 1995*

**Zusammenfassung:** Diese Arbeit bietet erstens einen Überblick über bestehende Systemtheorien und einige wichtige systemtheoretische Konstrukte. Zweitens will sie ein Gefühl dafür vermitteln, welche Erklärungspotentiale systemtheoretische Modellbildungen für die Kognitionswissenschaft haben. Dazu wird der Stand der Kunst bei der systemtheoretischen Modellierung intelligenzbezogener Phänomene in der Kognitionswissenschaft und einigen Nachbarwissenschaften ausführlich referiert und kommentiert. Und drittens will diese Arbeit auch ein Gefühl die praktischen und theoretischen Schwierigkeiten aufzeigen, welche mit der Verwendung systemtheoretischer Techniken einhergehen.

*erscheint 1996 in der Zeitschrift "Kognitionswissenschaft"*

## 1 Einführung

Systemtheoretische Modellbildungen erfreuen sich in vielen wissenschaftlichen Disziplinen rasch wachsender Beliebtheit. Sie gründet unter anderem in folgenden Punkten:

- Sehr verschiedene konkrete Systeme können mit einem universellen Instrumentarium untersucht werden. Diese Universalität gründet darin, daß sich moderne systemtheoretische Methoden einheitlich und umfassend *zeitlichen* Phänomenen widmen, oder anders ausgedrückt, daß Systeme als *dynamische* Systeme verstanden werden.
- Systemtheoretische Ansätze erhellen *komplexe* Systeme mit vielen Komponenten, Wechselbeziehungen und Wirkungsebenen.
- Exakte Analysen von Phänomenen der *Selbstorganisation* liefern neue Einsichten in teilweise seit der Antike rätselhafte Erscheinungen der Ordnungsbildung in der physikalischen, biologischen und sozialen Welt.

Intelligente Systeme -- Menschen, Tiere, vielleicht Roboter, vielleicht gewisse KI-Programme -- sind zweifellos dynamisch, komplex und selbstorganisierend. Es liegt also nahe, systemtheoretische Methoden in der Kognitionswissenschaft einzusetzen.

Gegenwärtig tauchen in der Kognitionswissenschaft die ersten systemtheoretischen Ansätze auf. Ihre methodische Bedeutung ist jedoch noch nicht geklärt. In der vorliegenden Arbeit<sup>1</sup> soll eine solche Klärung ansatzweise geleistet werden. Da systemtheoretische Techniken den meisten Lesern dieser Zeitschrift nur aus der Ferne bekannt sein dürften, biete ich im 2. Kapitel zunächst ein kleines Propädeutikum. Es besteht aus einer Übersicht über die verschiedenen existierenden Systemtheorien und einer informellen Beschreibung einiger zentraler Begriffsbildungen. Im zentralen 3. Kapitel wird dann referiert, wie systemtheoretische Methoden in der Kognitionswissenschaft und deren Nachbardisziplinen gegenwärtig verwendet werden. Dies Kapitel enthält auch zwei Exkurse zur Beziehung zwischen neuronalen Netzen und dynamischen Systemen sowie zur Natur reduktionistischer und nicht-reduktionistischer Erklärungen durch systemtheoretische Modelle. Das 4. Kapitel schließlich gibt eine Zusammenfassung und kritische Diskussion.

## 2 Ein kleines Propädeutikum zur Systemtheorie

Es gibt nicht "die" Systemtheorie, sondern viele mathematische, naturwissenschaftliche und gesellschaftswissenschaftliche Forschungslinien. Die verwendeten Systembegriffe sind teilweise unterschiedlich. Der gemeinsame Nenner besteht vielleicht darin, daß ein System etwas ist, das sich in der Zeit entwickelt und aus Teilsystemen oder Komponenten besteht, so daß das Ganze mehr als die Summe der Teile ist. Das ist nicht viel an Gemeinsamkeit.

Ein umfassender historischer Überblick wäre sehr wünschenswert. Leider scheint es ein solches Werk bislang nicht zu geben. Partielle historische Abrisse finden sich bei Abraham & Shaw (1992) (mathematikgeschichtlich), von Bertalanffy (1968) (20er und 30er Jahre, biologisch orientiert, Querbezüge zu anderen, vorzugsweise philosophischen und psychologischen Strömungen im deutschsprachigen Raum), Krohn, Küppers & Paslack (1987) (Ideengeschichte des Begriffs der Selbstorganisation), und Wiener (1961) (Dekade um den 2. Weltkrieg). Einen kurzen Abriß einiger Systemtheorien gibt auch von Weizsäcker (1987) in einem Aufsatz, der sich der Brückenfunktion der Systemtheorien zwischen Natur- und Geisteswissenschaften widmet. Eine umfassende historische Aufarbeitung der ingenieurwissenschaftlichen Systemtheorie liefert Wunsch (1985). S.J. Schmidt (1987b) vermittelt schließlich einen weitgehenden Überblick über die nicht-naturwissenschaftlichen Systemtheorien.

---

<sup>1</sup>Dieser Aufsatz ist eine überarbeitete und stark gekürzte Fassung von Jaeger (1995a).

Um für einen Einstieg in das Gebiet eine erste Orientierung zu geben, möchte ich nun einige ausgewählte Systemtheorien in Stichpunkten vorstellen.

**von Bertalanffy "Allgemeine Systemtheorie"** (von Bertalanffy 1968), schon in den 20er Jahren begonnen, ist als Methode der Biologie konzipiert. Erst nach dem 2. Weltkrieg kommt es zum Durchbruch mit der Gründung einer "Society for General System Theory" am Centre for Advanced Studies in Palo Alto 1954 (später "Society for General System Research"). Mathematisch werden Systeme durch Differentialgleichungen beschrieben. Im Gegensatz etwa zu Wiener (Kybernetik) oder Haken (Synergetik) entwickelt von Bertalanffy keine eigenen mathematischen Techniken. Von Bertalanffys Arbeiten muten philosophischer an als andere naturwissenschaftliche Systemtheorien. Diskutiert werden z.B. Fragen nach Finalität/Kausalität, Ganzheit/Gestalt, Relativität von Beschreibungskategorien. Die Lektüre von v. Bertalanffy lohnt auch heute noch, weil hier ein im Vergleich zu späteren Entwicklungen phänomenal besonders reichhaltiger und lebendiger Systembegriff vorliegt. Er umfaßt Aspekte wie funktionale und strukturelle Hierarchien, Wachstum, Automatisierung von Prozessen, Reflexbildung, progressive Segregation (d.h. Ausdifferenzierung eines Systems in immer unabhängiger Teilssysteme), Zentralisierung u.v.a.m. (alle in v. Bertalanffy 1986).

Die **Kybernetik** ist in der Zeit um den 2. Weltkrieg im engen intellektuellen Austausch von Wissenschaftlern aus Mathematik, Neurophysiologie und Psychiatrie, Regelungstechnik und anderen Gebieten entstanden (Beschreibung dieser Epoche in Wiener 1961<sup>2</sup>). Als zentrale Figuren kann man wohl den Mathematiker Norbert Wiener, den Ingenieur Julian H. Bigelow und den Neurologen Arturo Rosenblueth ansehen. Die Kybernetik wurde aber auch mit geformt durch von Neumann (Ergodentheorie, digitale Informationsverarbeitung), McCulloch und Pitts (erste mustererkennende neuronale Netzwerke, nach heutiger Terminologie), Lewin (Psychologie), Bateson und Mead (Anthropologie), sowie viele andere. Bei einer solch vielseitigen Ahnenschaft ist klar, daß es nicht *die* Kybernetik gibt. Tatsächlich wird bis in die heutige Zeit der Begriff oftmals fast nach Belieben für jedwede Forschungen gebraucht, die etwas mit Regelung, Informationsverarbeitung, biologischen Systemen oder anderen Facetten der Forschungen in der Gründerzeit zu tun haben. Theoretische Grundaufgaben sind Fragen der Zeitreihenvorhersage, der Stabilität und Oszillation in Systemen mit Feedback, und der Mustererkennung. Seit etwa den 60er Jahren sind die in der Kybernetik versammelten Themen wieder mehr in einzelne Disziplinen auseinandergetreten: die biologische Kybernetik, die Automatentheorie, die Informationstheorie, die ingenieurwissenschaftliche Kontrolltheorie, die Mustererkennung in der Informatik, und andere. Der Begriff "Kybernetik" hat jedoch überlebt, wenngleich auf Kosten einer nahezu beliebigen Verwendbarkeit.

Eine weitgehend selbständige **ingenieurwissenschaftliche Systemtheorie** (mit dem prominenten Teilgebiet der **Kontrolltheorie** bzw. deutscher ausgedrückt, der **Meß- und Regelungstechnik**) hat sich seit Beginn des Jahrhunderts in der Nachrichten- und Regelungstechnik entwickelt (geschichtlicher Abriß mit formaler Einführung zentraler Techniken: Wunsch 1985, Querschnitt von Einzeldarstellungen in Zadeh & Polak 1969, Lehr- und Handbuch Stengel 1986). Es ging zunächst um die Analyse der Eigenschaften von (elektrischen) Systemen bzw. um die Synthese solcher Systeme bei vorgegebenen Input-Output-Verhalten. Der Begriff "Systemtheorie" wird zuerst 1949 von Küpfmüller (referiert nach Wunsch 1985) verwendet. Küpfmüller behandelt komplexe elektrische Netzwerke, die er aus Teilsystemen zusammensetzt, welche formal nur durch ihr Input-Output-Verhalten beschrieben werden (Prinzip der "black box"). Die ingenieurwissenschaftliche Systemtheorie besitzt eine eigenständige, anwendungsorientierte Tradition, die von "akademischeren" Systemtheorien kaum rezipiert wird. Kontrolltheoretische Techniken (Filterung und Prädiktion stochastischer Prozesse) über-

schneiden sich vielfach mit der klassischen Kybernetik. Im Gegensatz zu anderen naturwissenschaftlichen Systemtheorien geht es nicht nur um die *Analyse* von Systemen, sondern mindestens ebenso sehr um den *Entwurf* von Systemen mit vorgegebenen gewünschten Eigenschaften, und um die *Kontrolle* von technischen Systemen.

**Zadehs "allgemeine Systemtheorie"** ist historisch aus der ingenieurwissenschaftlichen Tradition erwachsen. Inhaltlich ist sie aber eine selbständige, abstrakte, nahezu "reine" mathematische Theorie. Sie wurde in den frühen 60ern von L.A. Zadeh entwickelt (Einführung in Zadeh 1969). Systeme werden mengentheoretisch in einer solchen Allgemeinheit formalisiert, daß sich die in anderen mathematisierten Systemtheorien betrachteten Systeme (z.B. Differentialsysteme, Automaten) als Spezialfälle interpretieren lassen. Die mengentheoretische Spielart von Systemtheorie hat jedoch, wie es scheint, wenig Einfluß gewinnen können. Sie ist möglicherweise zu abstrakt, um noch für Anwendungen dienlich zu sein.

Die vom Neurobiologen Humberto R. Maturana und dessen Schüler Francisco J. Varela eingeführten **autopoietischen Systeme** sind ursprünglich ein Ansatz, Prinzipien lebendiger Systeme zu erklären. Maturana und Varela selbst bezogen bald Phänomene der Kognition und des Bewußtseins in ihre Untersuchungen ein (Einführung in Maturana & Varela 1984). Der Kernbegriff "Autopoiesis" ist schwer präzise zu fassen. Er betrifft die charakteristische Eigenschaft von Lebewesen, daß sie sich "*buchstäblich -- andauernd selbst erzeugen.*" (Maturana & Varela 1984 p.50f). Obwohl die Autoren mit konkreten physiologischen Beispielen nicht sparen, ist die Grundhaltung eher philosophisch und erkenntnistheoretisch als naturwissenschaftlich. Der Einfluß der Theorie autopoietischer Systeme in den Sozial- und Geisteswissenschaften, aber auch in der Psychiatrie und der Literaturwissenschaft, ist kaum zu überschätzen (Beispiele in S.J. Schmidt 1987b). Der Begriff "Systemtheorie" wird in solchen Gebieten oftmals synonym mit der von Maturana und Varela ausgehenden Tradition verwandt (z.B. Willke 1987). Die erkenntnistheoretische Grundthese, daß nämlich lebendige Systeme die äußere Wirklichkeit nicht intern abbilden, sondern je eine eigene, wesentlich durch das System bestimmte Wirklichkeit konstruieren, hat zu einer eigenen erkenntnistheoretischen Richtung in der Philosophie geführt, dem (radikalen) Konstruktivismus (Reader: S.J. Schmidt 1987a).

Auch **Evolutionstheorie** und **Populationsdynamik** (Wilson & Bossert 1973) können als Systemtheorien gelten. Die Populationsdynamik liefert einige Standardbeispiele dynamischer Systeme (Räuber-Beute-Systeme, Vermehrung unter begrenzten Ressourcen, vgl. Haken 1983 p.14). Die von Eigen & Schuster (1977f) im Hyperzyklus-Modell zur Entstehung des Lebens verwendeten mathematischen Methoden stammen aus der differentialgeometrischen Theorie dynamischer Systeme. Das Gebiet hat jedoch eine betont eigene Ausprägung. Es werden simultan sehr verschieden lange Zeitskalen berücksichtigt. Ökosysteme als (ein) Gegenstand der Untersuchung sind ausgeprägt offene Systeme. Das führt zu eigenen Modifikationen des Systembegriffs (Bemerkungen dazu in Zwölfer 1986). In der Evolutionstheorie geht es nicht nur um die Beschreibung zeitlicher Prozesse nach gegebenen Gesetzmäßigkeiten, sondern auch um die Mechanismen zeitlicher Veränderung selbst (Evolutionsmechanismen). Es wird sogar die Evolution von Evolutionsmechanismen erforscht (Wagner 1986). Einen Eindruck der Vielfalt evolutionstheoretischer Themen und Theorien vermitteln die Sammlung populärer Aufsätze von Gould (1979) und der methodologische Aufsatz von Wagner (1986).

**Physikalisch-mathematische Theorien der Selbstorganisation** sind wohl verantwortlich für den gegenwärtigen enormen Aufschwung systemtheoretischer Methoden auch außerhalb der Naturwissenschaften. Zu nennen sind hier vor allem die eben schon angesprochene Theorie des Hyperzyklus von Eigen und Schuster (1977), die Theorie der Selbstorganisation fern vom

thermodynamischen Gleichgewicht der "Brüsseler Schule" (Prigogine 1980), und die Synergetik von Haken (populäre Einführung in Haken & Wunderlich 1986, Standardreferenz Haken 1983). Die gemeinsame Fragestellung ist die Entstehung von Ordnung in Systemen, die aus (sehr) vielen gekoppelten Teilsystemen bestehen. Es zeigt sich, daß hinter dem intuitiven Begriff der Selbstorganisation eine große Phänomenvielfalt liegt. Die Synergetik und die Brüsseler Schule sind keine genau umrissenen formale Einzeltheorien, sondern Sammlungen analytischer Methoden zur Modellierung komplexer Systeme. Ein durchgängiges Thema sind Entwicklungspfade über Bifurkationen (Prigogine 1980, Kap. V, VI, Haken 1983 p. 264ff). Mit der Entdeckung und Erforschung des sog. deterministischen Chaos (populäre Einführungen in Davies 1987, Briggs & Peat 1989, Kriz 1992, etwas weniger populär in Abraham & Shaw 1992) sind seit etwa 15 Jahren systemtheoretische Ansätze weit über die exakten Wissenschaften hinaus als eine "Revolution" naturwissenschaftlichen Denkens bekannt geworden. Aus der Innenansicht stellt sich die "Chaostheorie" jedoch bescheidener (aber immer noch spannend genug) nur als eines von vielen Teilgebieten der modernen naturwissenschaftlichen Theoriebildung über dynamische Systeme dar (eine Bewertung aus naturwissenschaftlicher Sicht gibt Ford 1989).

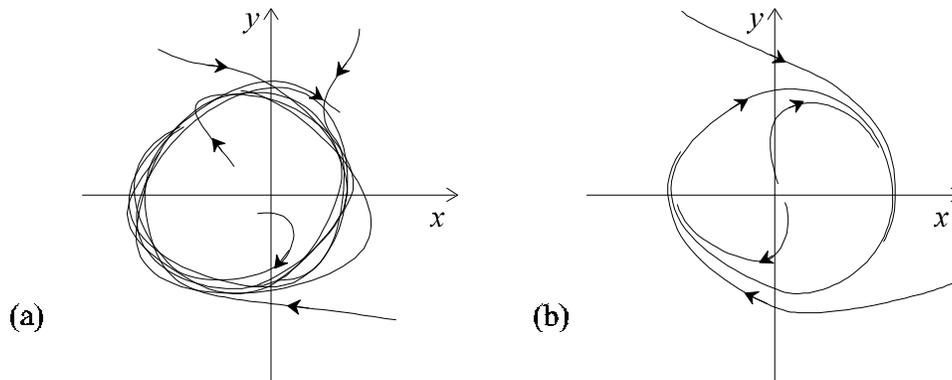
Die **mathematische Theorie dynamischer Systeme** hat sich in den letzten Dekaden zu einem außerordentlich fruchtbaren Zweig der mathematischen Forschung entwickelt. Man beschäftigt sich in einem vereinheitlichenden Rahmen sowohl mit Prozessen in kontinuierlichen Zustandsräumen (typischerweise durch Differentialgleichungen oder Abbildungen definiert) als auch in Symbolräumen (etwa durch Grammatiken oder Automaten bestimmt). Es geht, allgemein gesprochen, um die Analyse der *qualitativen* Eigenschaften solcher Prozesse, um Fragen der Stabilität, Komplexität, Vorhersagbarkeit und vieles andere. Die mathematische Systemtheorie ist ein ausgeprägt interdisziplinäres (bezogen auf Teilsparten innerhalb der Mathematik) Unternehmen, das von seinen Adepten gute Kenntnisse sehr verschiedener mathematischer "Sportarten" verlangt. Die atemberaubende Entwicklung der Systemtheorie zu einer ausgeprägten Teildisziplin der Mathematik in den letzten Dekaden sollte aber nicht vergessen machen, daß die mathematische Systemtheorie auf eine jahrhundertealte Tradition zurückblickt. Pendel, Bewegungen von Himmelskörpern, elektrische Schwingungen, Folgen von Würfelwürfen sind (neben Modellen des Bevölkerungswachstums) Urtypen dynamischer Systeme, die auch heute noch in keinem einschlägigen Lehrbuch ausgelassen werden, und sie stellen gleichzeitig den Kern der europäischen Mathematikgeschichte der Moderne dar. Eine sehr ausführliche, an diese Tradition anschließende, ebenso formelfreie wie dennoch exakte, überaus reich bebilderte Einführung bieten Abraham & Shaw (1992). Dieses ganz erstaunliche Werk (das schönste Fachbuch, das ich kenne) setzt keine Vorkenntnisse voraus! Als einführendes mathematisches Lehrbuch ist Arrowsmith & Place (1992) sehr zu empfehlen. Alle bis hier genannten Quellen betreffen ausschließlich Systeme mit kontinuierlichen Zustandsräumen. Katok & Hasselblatt (1995) haben ein umfassendes Handbuch erarbeitet, das auch den Bereich der symbolischen Prozesse abdeckt.

Hiermit möchte ich es gut sein lassen, auch wenn die Aufzählung noch unvollständig ist (es fehlen z.B. Thermodynamik und Informationstheorie, parallele, räumliche Systeme wie etwa zelluläre Automaten, Systemtheorien in der Wirtschaftslehre, u.a.m.). Nun möchte ich dynamische Systeme mit kontinuierlichen Zustandsräumen, im folgenden einfach dynamische Systeme (DS) genannt, etwas näher vorstellen. Nach Lage der einschlägigen Literatur sind für die Kognitionswissenschaft solche Systeme gegenwärtig die wichtigsten. Das im folgenden präsentierte Material betrifft einen einheitlichen Kernbestand an Begriffsbildungen, der von den naturwissenschaftlichen und mathematischen Systemtheorien geteilt wird. Als Beispiele verwende ich zumeist die "Klassiker" aus der Physik, um die Herkunft der Systemtheorie zu

akzentuieren. Bei der Darstellung verzichte ich auf mathematische Formeln. Etwas strengere Einführungen finden sich bei Kelso et al. (1993) und Jaeger (1995a).

Zunächst einige handwerkliche Bemerkungen dazu, wie man überhaupt ein DS formal "hinschreibt". Ein dynamisches System besteht aus einem kontinuierlichen *Zustandsraum* (synonym: *Phasenraum*), dem eine Dynamik aufgeprägt ist, d.h. eine Gesetzmäßigkeit, welche festlegt, wie im Laufe der Zeit die Zustände aufeinander folgen. Ein solcher Zustandsraum ist durch reellwertige Dimensionen  $x_1, \dots, x_n$  bestimmt. Ein Zustand ist also ein  $n$ -dimensionaler Vektor  $(x_1, \dots, x_n)$ .

Betrachtet man ein DS als Modell für ein reales System, so repräsentiert jede Dimension eine "Meßgröße", eine Observable. Man nennt sie auch *Systemvariable*, *Systemgrößen*, *Systemparameter*, oder noch anders. Ein Zustandsvektor  $(x_1, \dots, x_n)$  gibt also an, welche numerischen Werte zu einem gewissen Zeitpunkt die beobachteten Systemvariable einnehmen.



**Abb. 1:** Ein empirisch gewonnenes, zweidimensionales Phasenportrait (a) und eine formale Rekonstruktion davon (b). Pfeile zeigen die Zeitrichtung der Trajektorien an.

Die Systemvariable verändern sich im Laufe der Zeit. Ich möchte mich hier auf den Fall beschränken, daß diese Veränderung *kontinuierlich* erfolgt (zu den genauso wichtigen, mathematisch eng verwandten Dynamiken mit *diskreten* Zeitschritten vgl. Jaeger 1995b). Eine kontinuierliche Veränderung der Systemvariable führt dazu, daß das System eine zusammenhängende Systemgeschichte, eine "Spur in der Zeit", oder wie der Fachausdruck lautet, eine *Trajektorie* verfolgt. Trägt man eine Trajektorie (oder mehrere "typische" davon) in einem Diagramm auf, erhält man ein *Phasenportrait* (synonym: *Phasenbild*). Abbildung 1 zeigt zwei solche Phasenportraits. Abb. 1a mag etwa aus der empirischen Aufzeichnung typischer Spuren des Zeigefingers einer Versuchsperson entstammen, der man den Auftrag gegeben hat, von verschiedenen Startorten aus eine Kreisbewegung im Uhrzeigersinn auszuführen. Abb. 1b zeigt eine Version von 1a, wie sie einer "sauber" gerechneten Simulation desselben Vorganges entspringen würde, die auf einer differentialgeometrischen Rekonstruktion des Systems beruht. Die Zeit  $t$  ist in Phasenportraits nicht als eigene Achse berücksichtigt, sondern implizit im Verlauf der Trajektorien eingefangen.

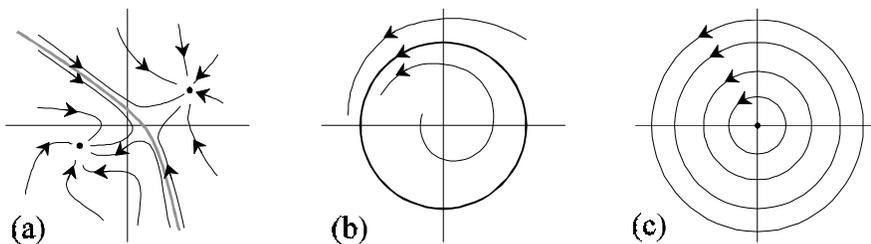
Mathematisch werden DS meistens durch Differentialgleichungen beschrieben, worauf ich hier nicht weiter eingehe. Im Unterschied zu empirischen Phasenportraits können sich Trajektorien in solchermaßen mathematisch modellierten DS nicht kreuzen (vgl. Abb. 1b).

Ein zentrales Phänomen in DS sind *Attraktoren*. Intuitiv ist ein Attraktor ein Gebiet im Phasenraum, das benachbarte Trajektorien "anzieht", indem sie sich diesem Gebiet asymptotisch nähern. Der einfachste Fall sind punktförmige Attraktoren, sog. *Fixpunktattraktoren*. Abb. 2a zeigt ein Phasenportrait mit zwei Fixpunktattraktoren. Trajektorien, die im "Anzie-

hungsbereich" (dem sog. *Attraktionsbecken*) des Fixpunktattraktors liegen, entwickeln sich auf diesen zu und nähern sich ihm asymptotisch beliebig nahe an. Wenn das System zufällig gerade in einem solchen Punkt gestartet wird, verharrt es einfach darin: d.h. der Fixpunktattraktor selbst ist eine (punktförmige) Trajektorie. Fixpunkt-Attraktoren sind in empirischen Systemen sehr häufig, sie entsprechen intuitiv gesehen "Ruhezuständen". Das klassische Beispiel ist ein durch Reibung gedämpftes Pendel. Ruhezustände sind *stabil*. Das bedeutet, daß kleine "Störungen" des Systems (etwa wenn eine Fliege gegen das Pendel fliegt und es etwas aus seiner Ruhelage auslenkt) nach kurzer Zeit wieder ausgeglichen werden, indem die Trajektorie zum Fixpunktattraktor zurückkehrt. Bei andauernden kleinen Störungen (etwa durch Stöße von Luftmolekülen) bedeutet die Stabilität, daß sich das System mit seinen Zuständen in einer kleinen Umgebung des Ruhezustandes aufhält.

In Abb. 2b (und in Abb. 1) finden wir den zweiteinfachsten Fall, einen *zyklischen Attraktor*. Der Attraktor selbst ist eine geschlossene, zyklische Trajektorie. Man sieht, daß Trajektorien aus der Umgebung des Attraktors sich bei  $t \rightarrow \infty$  beliebig nahe spiralförmig diesem nähern. Zyklische Attraktoren sind, intuitiv gesehen, "stabile Schwingungen". Das klassische Beispiel liefert das angeregte Pendel, z.B. das Pendel einer Uhr. Auch hier heißt Stabilität wieder, daß bei kleinen Störungen der Attraktor von der Systemtrajektorie asymptotisch wieder eingenommen wird. Zyklische Attraktoren sind ein sehr häufiges Phänomen in dynamischen Systemen. Man muß sich beim Erfinden von Systemgleichungen nicht besonders anstrengen, um einzelne Gebiete im Phasenraum mit stabilen Schwingungen auszustatten. Diese Ubiquität von stabilen periodischen Aktivitätsmustern kann in ihrer Bedeutung für die Welt, wie wir sie kennen (und die uns hervorgebracht hat), nicht überschätzt werden. Gäbe es nämlich gar keine Attraktoren, so gäbe es, grob gesprochen, überhaupt keine stabilen Phänomene in der physikalischen Realität. Gäbe es jedoch nur Fixpunktattraktoren, so würde die Realität in einem gigantischen, finalen Ruhezustand ersterben. Erst zyklische Attraktoren eröffnen die Möglichkeit, daß einerseits sich etwas "tut", andererseits aber stabile, wiedererkennbare, klassifizierbare Phänomene in die Welt kommen -- daß nicht alles nur in ungegliederter Irregularität vergeht. Es ist darum geradezu denknotwendig, daß in der Physik Elementarteilchen mathematisch als stabile Schwingungen beschrieben werden. Der Substanzbegriff selbst ist also in heutiger Sicht mit zyklischen Attraktoren eng verbunden. Die Bedeutung von Periodizitäten und Rhythmen für das Gebiet der Kognitionswissenschaft wird im folgenden Kapitel ausführlich gewürdigt.

Zum Kontrast zeigt Abb. 2c ein Phasenportrait, in dem die Trajektorien parallel nebeneinander herlaufen. Nichts nähert sich irgendetwas anderem - hier gibt es keinen Attraktor!



**Abb. 2:** Die beiden einfachsten Attraktor-Typen. (a) Fixpunkt-Attraktor (zwei Stück; die schattierte Linie markiert die Grenze zwischen den beiden zugehörigen Attraktionsbecken), (b) zyklischer Attraktor, (c) kein Attraktor in diesem Phasenportrait!

Es gibt mit zyklischen Attraktor verwandte Erscheinungen, die etwas komplizierter aussehen als dieser. Sie entstehen, wenn sich mehrere solcher Attraktoren überlagern. Stehen die Periodenlängen der dabei beteiligten Grenzzykel nicht in einem rationalen Verhältnis zueinander,

der, so ergibt sich keine zyklisch geschlossene Trajektorie mehr. Vielmehr liegt der dann *quasi-periodisch* genannte Attraktor dicht auf einem Torus (vgl. Haken 1983, 28f).

Neben Fixpunkt- und zyklischen Attraktoren gibt es noch eine dritte Klasse. Sie wird erst seit den 60er Jahren systematisch untersucht, obwohl schon zu Beginn des Jahrhunderts verschiedentlich Mathematiker über dieses Phänomen stolperten, sich aber, letztlich mangels Computergraphik, im wahrsten Sinne des Wortes kein Bild davon machen konnten. Es handelt sich um die *chaotischen* Attraktoren (zunehmend selten auch *seltsame* - engl. *strange* - Attraktoren genannt). Sie sehen auf den ersten Blick wie sehr komplizierte zyklische Attraktoren aus. Abb. 3 zeigt den einfachsten bekannten Vertreter seiner Gattung, den von Otto Rössler (1968) gefundene, nach ihm benannten chaotische Attraktor.

**Abb. 3:** Eine Ausprägung des Rössler-Attraktors (aus Nicolis 1989, p. 333)

In Abb. 3 ist nur der Attraktor selber skizziert, nicht die sich an ihn "anschmiegenden" Trajektorien aus seiner Nachbarschaft. Tatsächlich sind chaotische Attraktoren noch kompliziertere Gebilde als die Abbildung erahnen läßt. Sie bestehen nämlich aus unendlich vielen Trajektorien, die in einer *fraktalen* Struktur angeordnet sind (Einführung bei Kriz 1992, einfach zu lesender Klassiker Mandelbrot 1977). Grob gesagt, sind sie durch *Selbstähnlichkeit* über alle Größenordnungen hinweg charakterisiert. Das besagt folgendes. Wenn man an einer Stelle in Abb. 3 einen Querschnitt durch den Attraktor legen würde, könnte man das entstehende Punktemuster immer wieder ausschnittsvergrößern, ohne daß sich wesentliche Änderungen ergäben.

Eine zentrale Eigenschaft chaotischer Attraktoren ist deren Unvorhersehbarkeit. Wenn man zwei Trajektorien aus dem Rössler-Attraktor verfolgt, die zu einem gewissen willkürlichen Startzeitpunkt beliebig dicht nebeneinanderliegen, haben die beiden damit fixierten Systementwicklungen schon nach kurzer Zeit nichts mehr miteinander gemein. Diese Divergenz von Systemgeschichten führt angesichts von prinzipiellen (wenn auch noch so kleinen) Ungenauigkeiten bei der Messung von initialen Systemzuständen dazu, daß man Systeme in chaotischen Attraktorzuständen nicht vorhersehen kann. Umgekehrt lassen sich Systeme, die auf den ersten Blick als vollkommen erratisch erscheinen, oft als Systeme in chaotischen Attraktorzu-

ständen deuten -- sie weisen eine "verborgene Ordnung" auf und sind entgegen dem äußeren Anschein nur von wenigen Systemvariablen und einem oft einfachen, deterministischen Gesetz bestimmt.

Die (Wieder-)Entdeckung chaotischer Attraktoren, fraktaler Strukturen und damit verbundener Erscheinungen hat etwa ab Mitte der 60er Jahre zu einem ungeheuren Aufschwung in der Erforschung dynamischer Systeme geführt. Die landläufige, inzwischen zum geistigen Allgemeinut gehörende Bezeichnung "Chaostheorie" (die es als eigenständige Theorie nicht gibt) reflektiert diese Entwicklung. Chaotische Systeme sind zwar deterministisch, jedoch verhalten sie sich in puncto Vorhersagbarkeit in vielem wie stochastische Systeme. Dies führt zum Begriff des "deterministischen Chaos", der geeignet erscheint, einige alte erkenntnistheoretische Fragen neu zu interpretieren. Heute sind viele Naturwissenschaftler schon generell geneigt, hinter jedem äußerlich als zufällig erscheinenden Prozeß eine verborgene chaotische Ordnung zu sehen. Eine Warnung davor, die Rolle von "echtem" Zufall geringzuschätzen, gibt Millonas (1994).

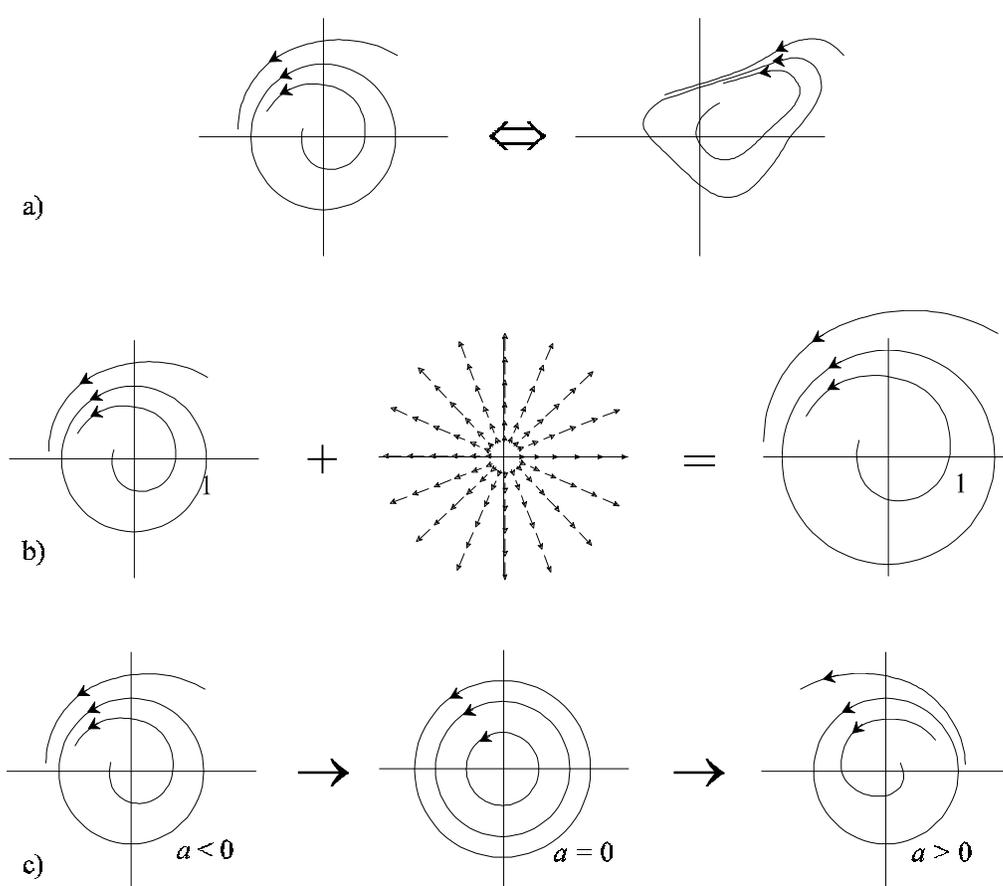
Genauso wichtig für ein Verständnis dynamischer Systeme wie Attraktoren sind *Bifurkationen*. Sie betreffen das Phänomen, daß sich ein System plötzlich qualitativ verändert, einen "Umbruch" durchmacht. So ist es zum Beispiel aus systemtheoretischer Sicht eine Bifurkation, wenn man mit dem Arm gegen ein Hindernis drückt, dabei die Kraft zunehmend vergrößert, bis plötzlich der Arm zu zittern beginnt. In diesem Moment verändert sich ein Fixpunktattraktor (der die stabile Armhaltung beschreibt) plötzlich zu einem zyklischen Attraktor (der die Zitterbewegung beschreibt). Um den Bifurkationsbegriff genauer zu erklären, muß ich etwas ausholen und zuvor die Begriffe der topologischen Äquivalenz, der strukturellen Stabilität und der Kontrollparameter erläutern.

Zwei dynamische Systeme heißen *topologisch äquivalent*, wenn das Phasenportrait von  $S_1$  eins-zu-eins so auf das Phasenportrait von  $S_2$  abgebildet werden kann, daß Trajektorien auf Trajektorien abgebildet werden, wobei deren Orientierung erhalten bleibt. Man kann sich das intuitiv so vorstellen, daß das Phasenportrait von  $S_1$  auf ein Blatt aus Gummi aufgemalt wird, und dieses dann verzerrt wird, um das Phasenportrait von  $S_2$  zu erhalten. So sind etwa die beiden Phasenportraits in Abb. 4a topologisch äquivalent. Etwas appellativ gesprochen, sind zwei DS topologisch äquivalent, wenn sie "qualitativ gleichartig" sind.

Daß ein dynamisches System *strukturell stabil* ist, heißt, daß es mit sich selbst topologisch äquivalent bleibt, wenn es durch beliebige, aber kleine Störungen modifiziert wird. Die meisten dynamischen Systeme sind meistens strukturell stabil - in einer Welt voller Störeinflüsse hätte es beunruhigende Folgen, wenn Systeme nicht meistens ihre qualitativen Typen bewahrten! Abbildung 4b illustriert das Prinzip. Die kleinen nach außen strebenden "Kräftepfeile" im mittleren Diagramm sollen eine vom Betrag her kleine Störung des Systems darstellen. Diese Störung bewirkt lediglich, daß das Phasenportrait gleichsam etwas größer wird. Es bleibt aber offensichtlich in seinem qualitativen Typ erhalten.

Manche Systeme sind nicht strukturell stabil. So würde etwa das System in Abb. 2c auf die meisten kleinen Störungen damit reagieren, daß die zyklisch geschlossenen Trajektorien "sich selbst nicht mehr trafen", also zu Spiralen würden. Ein Phasenportrait mit Spiralen kann aber nicht eins-zu-eins auf eines der Art von Abb. 2c projiziert werden -- daher hätte also das System aus Abb. 2c seinen qualitativen Typ verändert. Es ist nicht strukturell stabil.

*Kontrollparameter* schließlich betreffen das Faktum, daß dynamische Systeme in ihrem Verhalten von außen durch "Steuergrößen" beeinflusst werden können. So wird etwa ein elektrischer Schwingkreis sich je nach angelegter Spannung oder nach Einstellwert des Drehkondensators anders verhalten. Der schon erwähnte Arm zeigt ein von der aufgewendeten Kraft abhängiges Verhalten. Spannung, Kondensatoreinstellung und Kraft sind Beispiele für Kontrollparameter.



**Abb.4:** Topologische Äquivalenz (a), strukturelle Stabilität (b), und eine Bifurkation (c).

Nun können wir die Grundsituation, in der Bifurkationen auftreten, folgendermaßen beschreiben. Man nehme ein System, das durch einen Kontrollparameter  $a$  beeinflusst werden kann. Wir halten  $a$  fest und nehmen ein Phasenportrait auf. Dann verändern wir  $a$  ein wenig (wir lassen seinen Wert z.B. etwas wachsen) und nehmen erneut ein Phasenportrait auf. Dank der strukturellen Stabilität, die das System beim ersten Mal (höchstwahrscheinlich) besaß, wird das zweite Phasenportrait zum ersten topologisch äquivalent sein, wenn auch etwas "verformt". Wir wiederholen diese kleinen Veränderungsschritte von  $a$  und werden feststellen, daß das System über einen gewissen Bereich qualitativ stabil bleibt. Zumeist können wir dann aber plötzlich einen Umschlag zu einem qualitativ neuartigen Phasenportrait sehen, wenn  $a$  über einen bestimmten Wert hinaus wächst. Jenseits dieses Wertes bleibt bei weiteren kleinen Veränderungen das System wieder strukturell stabil. Und wenn wir den Kontrollparameter *genau* auf den kritischen Wert einstellen, werden wir ein strukturell instabiles Phasenportrait sehen können. Es markiert den einzelnen Bifurkationspunkt, der zwischen zwei Stabilitätsbereichen liegt. In Abbildung 4c findet sich ein Beispiel, wo im Bereich  $a < 0$  das System qualitativ durch einen zyklischen Attraktor charakterisiert ist, und für  $a > 0$  durch das genaue Gegenteil, nämlich einen zyklischen *Repellor* (d.h. eine zyklisch geschlossen Trajektorie, die benachbarte Trajektorien "abstößt"). Für den kritischen Wert  $a = 0$  ergibt sich das inzwischen vertraute, strukturell instabile Phasenportrait konzentrischer Zykel.

Diese Skizze einer Bifurkation betrifft den "typischen" Fall, der für Anwendungen in der Kognitionswissenschaft wichtig ist. Er ist durch den Dreischritt Stabilitätsbereich  $\rightarrow$  Bifurkationspunkt  $\rightarrow$  qualitativ neuartiger Stabilitätsbereich gekennzeichnet. Der Bifurkationsbegriff wird allerdings lose auch für einige andere Phänomene des qualitativen Umschlags verwendet.

Genauso wie Attraktoren sind Bifurkationen ein ubiquitäres Phänomen in dynamischen Systemen. Wir könnten uns eine Welt ohne qualitative Umsprünge zwischen mehr oder weniger stabilen Systemzuständen schlechthin kaum vorstellen: eine qualitative Entwicklung ist immer das Erscheinen von etwas "ganz Neuem" -- *qualitative Entwicklungen können sich nicht graduell vollziehen!* Der alte Gedanke vom Umschlag von Quantität in Qualität bekommt eine präzise Bedeutung.

Die Bifurkationstheorie ist eines der folgenreichsten und auch anspruchsvollsten Teilgebiete der mathematischen und naturwissenschaftlichen Systemtheorie. Es sind inzwischen eine ganze Reihe verschiedener Typen von Bifurkationen eingehend analysiert worden. Darüber hinaus interessiert man sich auch für ganze "Kaskaden" von Bifurkationen, die unter geeigneten Umständen entstehen, wenn ein Kontrollparameter immer weiter verändert wird, so daß sich viele Bifurkationen aneinanderreihen. Über solche Bifurkationskaskaden kann man z.B. beschreiben, wie ein System von einem einfachen Fixpunktattraktor über (unendlich) viele Zwischenstufen zu einem chaotischen Verhalten gebracht werden kann. Auch dies scheinbar ausgefallene Szenario entpuppt sich bei näherem Hinsehen als nahezu allgegenwärtig in der physikalischen Realität!

Zum Abschluß dieses Kapitels soll noch kurz auf das *Versklavungsprinzip* hingewiesen werden. Es ist ein originärer Beitrag der Synergetik (Haken 1983, p. 35f, Kapitel 7). Kurz gefaßt, besagt es, daß gewisse (wenige) Systemvariable die restlichen (vielen) dominieren, in dem Sinne, daß sich aus der Kenntnis der Dynamik der ersteren die Dynamik der letzteren ableiten läßt. Die wenigen, dominierenden Variablen werden auch *Ordnungsparameter* genannt (nicht zu verwechseln mit Kontrollparametern), und man sagt, sie *versklaven* die restlichen Parameter. Das Standardbeispiel ist der Laser. Wenn er "lasert", schwingen intuitiv gesprochen alle seine atomaren Mikro-Oszillatoren im Gleichtakt und in einer Richtung -- zur Beschreibung des Systems reicht also eine einzige Systemvariable aus. Befindet sich der Laser dagegen im "unterkritischen" Zustand, so schwingen alle Mikro-Oszillatoren im wesentlichen unabhängig voneinander und müssen für eine Kenntnis des Systems alle einzeln beschrieben werden.

Damit sei das kleine Propädeutikum beendet.

### **3 Dynamische Systeme in der Modellierung kognitiver Prozesse**

Dies Kapitel bietet einen ansatzweisen Überblick über den Stand der Kunst bei der Verwendung dynamischer Systeme zur Modellierung kognitiver Prozesse. Die Kognitionswissenschaft im engeren Sinne ist -- wie die direkten Nachbargebiete der kognitiven Psychologie, der KI und der Linguistik -- traditionell eher logikorientierten Methoden auf "höheren" Ebenen der Kognition verbunden. Dynamische Systeme werden hier bislang noch selten eingesetzt. Häufiger finden sie in etwas weniger direkt benachbarten Gebieten Verwendung, etwa in der Neurowissenschaft und dem Konnektionismus. In einigen eher mittelbar benachbarten Disziplinen schließlich gehören systemtheoretische Methoden zum Standardrepertoire: in der Ethologie, der Robotik, der Artificial-Life-Forschung, sowie gewissen Sparten der theoretischen Physik, die sich (von der Kognitionswissenschaft unbemerkt) mit Grundfragen symbolischer und neuronaler Informationsverarbeitung befassen. Meiner Meinung nach kann die Kognitionswissenschaften einen großen Gewinn daraus ziehen, wenn sie all diese Entwicklungen zur Kenntnis nimmt und einiges daraus für ihr eigenes Methodenarsenal adaptiert. Deshalb werde ich mich in diesem Kapitel nicht auf ein Referieren des Standes der Kunst innerhalb der Kognitionswissenschaft im engeren Sinn beschränken. Vielmehr möchte ich aus allen genannten Disziplinen relevantes Material zusammentragen und möglichst zu einem Gesamtbild verbinden. Dabei verste-

he ich systemtheoretische Methoden nicht als Ersatz oder Alternative zu den logisch-symbolischen, sondern als Ergänzung.

Das Kapitel ist in vier Abschnitte gegliedert. Zuerst wird die gegenwärtige Bedeutung systemtheoretischer Methoden in den einzelnen Disziplinen knapp umrissen (3.1). Den Hauptteil nimmt die Darstellung einzelner inhaltlicher Aspekte intelligenter Informationsverarbeitung ein (3.2). In zwei ergänzenden Abschnitten finden sich Bemerkungen über die Beziehung zwischen neuronalen Netzen und dynamischen Systemen (3.3) sowie über die (nicht-)reduktionistische Natur systemtheoretischer Erklärungen (3.4).

### **3.1 Systemtheoretische Methoden in den einzelnen Fachdisziplinen: eine knappe Orientierung**

In der **Kognitionswissenschaft** und der **Psychologie** spielen systemtheoretische Methoden besonders bei der Erklärung motorischen Verhaltens seit Beginn der 80er Jahre eine zunehmend wichtige Rolle. Der Impetus ist dabei wohl von der Synergetik in die Psychologie hineingetragen worden (cf. Referenzen in Schöner et al. 1986). Die formale Modellierung motorischer Systeme als dynamische Systeme ist inzwischen etabliert, und die hier erzielten Erkenntnisse inspirieren die Übertragung der rezipierten Methoden in andere Teilgebiete der Kognitionswissenschaft und Psychologie (Sammelband: Smith & Thelen 1993a). Neuerdings werden dynamische Systeme sogar als ein fundamentales Paradigma zur Modellierung kognitiver Prozesse vorgeschlagen (van Gelder & Port 1994b, van Gelder 1995). Außerdem werden formale Methoden dynamischer Systeme durch konnektionistische Techniken in das Gebiet importiert.

In der klassischen **künstlichen Intelligenz** sind systemtheoretische Methoden noch kaum in Erscheinung getreten. Mir sind nur zwei Fälle bekannt. Hasida (1994) implementiert eine Kontrollstrategie für *fuzzy*-Prolog-Programme, welche auf einem Energieminimierungsprinzip in einem kontinuierlichen dynamischen System beruht. Und ich selber (Jaeger 1994a,b,c,d) entwickle einen Typ diskreter, selbstorganisierender sog. *dynamischer Symbolsysteme* als formales Modell intelligenter Agenten. Eine Sonderrolle nehmen die Arbeiten von Sacks (1990, 1991) ein. In einem *qualitative reasoning*-Kontext beschreibt er, wie mit klassischen KI-Schlußfolgerungsmethoden *über* dynamische Systeme räsoniert wird.

In der **Linguistik** besitzen strukturalistische Ansätze eine gewisse Geistesverwandtschaft zu systemtheoretischen Methoden. Jedoch sieht die strukturalistische Tradition die Sprache zwar als System, aber nicht als dynamisches System. Hierzu gibt es eine Gegenbewegung, die den dynamischen Charakter sprachlicher Prozesse betont (Sammlung in Rieger 1985), ohne allerdings eigentlich systemtheoretische Methoden einzuführen. Es gibt jedoch vereinzelte Versuche, gerade dies zu tun (Köhler 1987, Rickheit & Strohner 1992, Tucker & Hirsh-Pasek 1993). Hierbei handelt es sich um Plädoyers, welche mit eher intuitiven Argumenten systemtheoretische Methoden für die Linguistik propagieren. Wildgen (1985) entwirft eine umfassende, mathematisch in der Katastrophentheorie (einer klassischen Bifurkationstheorie) gegründete Theorie zur Satzsemantik. Strengere systemtheoretische Techniken gelangen auch gelegentlich auf dem Umweg über den Konnektionismus in die Linguistik (z.B. Smolensky 1986, Legendre et al. 1990a,b).

Die Erforschung **autonomer adaptiver Agenten** und der Robotik-Zweig der **Artificial-Life-Forschung** entwickelt sich zu einem eigenständigen Gebiet zwischen Biologie, Kognitionswissenschaft, KI und Robotik. Systemtheoretische Ansätze kommen hier in dreifacher Weise zum

Tragen. Erstens sind jüngsthin Vorschläge gemacht worden, Agent-Umgebungs-Systeme formal durch gekoppelte, kontinuierliche dynamische Systeme zu modellieren (Smithers 1994, Beer 1995). Diese Formalisierungen können als vorläufige Verwirklichung von programmatischen Äußerungen gelten, die von Proponenten des *Situated Action*-Paradigmas gemacht worden sind (z.B. Clancey 1993, Greeno & Moore 1993). Zweitens werden Techniken der Kontrolltheorie durch die Robotik und neuronale Netze in das Gebiet importiert (z.B. Weidemann & Pfeiffer 1994, Steinkühler & Cruse 1994). Drittens werden durch die Thematik Adaptivität/Lernen in Einzelfällen systemtheoretische Techniken assimiliert (z.B. Kaelbling 1991, Steels 1994b). Insgesamt sind dem programmatischen Anspruch nach systemtheoretische Methoden etabliert, wenn auch die konkrete Ausgestaltung erst begonnen hat.

In der **Neurophysik**, **Biokybernetik** und der **Neuronale-Netze-Forschung** bilden Methoden der Kybernetik, Kontrolltheorie, der naturwissenschaftlichen und der mathematischen Theorie dynamischer Systeme die Basis der meisten formalen Modelle. Das mathematische Niveau ist hoch. Es gibt mehrere einschlägige Zeitschriften und eine Vielzahl spezialisierter Forschungsrichtungen. Eine Übersicht kann hier nicht versucht werden. Arbeiten aus dem Gebiet werden in 3.2 reichlich zitiert.

In der **Robotik** gehören Methoden der Kontrolltheorie naturgemäß zum Standardrepertoire ingenieursmäßiger Techniken der Bewegungskontrolle und der Signalverarbeitung (z.B. in reiner Form bei Triggs 1994, in Kombination mit neuronalen Netzwerken bei Schweitzer & Wen 1994). In den letzten Jahren sind auch Prinzipien der Selbstorganisation ausgenutzt worden, teilweise in Anlehnung an biologische Vorbilder (z.B. Gaudiano & Grossberg 1991, Cruse et al. 1994).

In der **theoretischen Physik** und der **mathematischen Ergodentheorie** werden auf sehr allgemeinem Niveau symbolische Prozesse untersucht, d.h. Mechanismen, welche Symbolfolgen generieren. Dabei wird von Physikern gerne die Verbindung zu kognitiven Prozessen hergestellt (z.B. J.S. Nicolis & Tsuda 1989, Ebeling & G. Nicolis 1992). Es existieren hier hochentwickelte statistische Analysetechniken für symbolische Zeitserien, die z.B. genaue Rückschlüsse über erforderliche Gedächtnisleistungen zulassen. Crutchfield (1992) entwickelt eine informationstheoretische Semantiktheorie für dynamische Prozesse und deren formale Beschreibungen, welche in vielem über die in der Kognitionswissenschaft vertraute Modelltheorie hinausweist.

## 3.2 Systemtheoretische Beiträge zum Verständnis intelligenter Systeme

In diesem Abschnitt möchte ich einen lebendigen Eindruck von den Möglichkeiten dynamischer Systeme bei der Modellierung kognitiver Prozesse vermitteln. Zunächst behandle ich mit Periodizitäten und Rhythmen (3.2.1) einen Bereich, wo "Dynamik" vielleicht am auffälligsten in Erscheinung tritt. Dann wende ich mich mit Fragen der Entwicklung, Differenzierung und des Lernens (3.2.2) längerfristigen Prozessen zu. Anschließend stelle ich systemtheoretische Beiträge zur Konzeptmodellierung (3.2.3), zu Agent-Umwelt-Systemen (3.2.4), und zu Fragen der Semantik (3.2.5) vor.

Um bei den zitierten Arbeiten die Zuordnung zu Fachgebieten zu ermöglichen, gebe ich den disziplinären Hintergrund der Autoren jeweils durch Kürzel mit an (subjektive Einschätzung -- ich bitte die Betroffenen um Verzeihung). Es bedeuten: *AA* = autonome Agenten, *Bio* = Biokybernetik und Neurowissenschaft, *Inf* = Informatik, *KI* = Künstliche Intelligenz, *Ling* = Linguistik, *Kog* = Kognitionswissenschaft, *Math* = Mathematik, *NN* = Neuronale-Netze-For-

schung, *Phil* = Philosophie, *Phys* = Physik, *Psych* = Psychologie, *Rob* = Robotik und ingenieurmäßige Mustererkennung, *Soz* = Sozial- und Wirtschaftswissenschaft.

### 3.2.1 Zeitmuster

Man könnte sagen, nur im Tod fließt die Zeit gleichförmig. Lebendiges Geschehen bildet Muster in der Zeit -- Schwingungen und Rhythmen, die in einer unübersehbaren Vielfalt auftreten, von präzisen Oszillationen bis zu unvorhersehbaren Orbits im Chaos. Dynamische Systeme stellen die formalen Mittel bereit, um diese Phänomene zu beschreiben und Universalien aufzudecken. F.D. Abrahams (*Psych* 1990) skizziert in einer großangelegten Einführungsarbeit, wie in den verschiedensten Teildisziplinen der Psychologie (von der Jung'schen Archetypenlehre über Gruppenpsychologie, Attribuierungstheorien und behavioristische Lerntheorien bis zu neurologischen Ansätzen) einfache Zeitmuster in universaler Weise auftreten und mit Standardtechniken der Systemtheorie beschrieben werden können. Etwas nüchterner stellen van Gelder & Port (*Kog* 1994b) die Grundtatsache der Zeitlichkeit kognitiver Prozesse als ein zwingendes Motiv für die Verwendung dynamischer Systeme in der Kognitionswissenschaft heraus. Ich möchte in diesem Abschnitt in einer Art "Wahrnehmungsschulung" die Sinne schärfen für die reiche Phänomenologie des Zeitlichen im mentalen und kognitiven Bereich.

Am deutlichsten tritt die Phänomenologie der Zeitlichkeit wohl dort in den Vordergrund, wo es direkt um die **Wahrnehmung und Reproduktion zeitlicher Muster** geht. Hier gibt es noch nicht viele, dafür aber schöne Arbeiten zu referieren. Port et al. (*Kog, NN* 1994) finden durch experimentelle und formale Untersuchungen, daß Menschen auditive Zeitmuster nicht anhand einer absoluten internen Referenzzeit analysieren, sondern durch mehrere spezialisierte Verarbeitungstechniken, die u.a. periodische Regelmäßigkeiten im Signal selbst für Referenzzwecke ausnutzen. Subtile Phänomene der Wahrnehmung und Prädiktion von Rhythmen modellieren Large & Kolen (*Psych* 1994), indem sie neuronale Module als gekoppelte, phasen- und frequenzschließende Oszillatoren spezifizieren. Diese sehr klar geschriebene Arbeit kann auch als Einführung in die neuronale Modellierung der Verarbeitung von Rhythmen, und das Gebiet dynamischer Systeme aus gekoppelten Oszillatoren gelesen werden.

Gekoppelte Oszillatoren dienen auch als mathematisches Modell in einer inzwischen wohl klassisch zu nennenden Studienserie zur Analyse periodischer Handbewegungen (Schöner et al. *Phys Psych* 1986, einführende Beschreibung in Kriz *Psych* 1992, 150ff). Versuchspersonen wurden aufgefordert, mit den Händen gegenläufige periodische Bewegungen vorgegebener Frequenz auszuführen. Bei Steigerung der Frequenz kommt es zu einem unwillkürlichen Umschlag in eine gleichläufige Bewegung. Die empirisch gefundenen stochastischen Charakteristika des Bewegungsumschlages konnten im Detail erklärt werden. Der Bewegungsumschlag selbst ist eine Bifurkation zwischen verschiedenen Attraktoren. Diese Arbeiten haben viele Nachfolgeuntersuchungen ausgelöst (vgl. Smith & Thelen 1993b).

Wenn man die Verarbeitung zeitlicher Muster in kognitiven Systemen untersucht, stellt sich die Frage nach der **Speicherung zeitlicher Information**. Schon Lashley (*Bio* 1951) hat in einem immer noch aktuell anmutenden Aufsatz die Frage aufgeworfen, wie es zur Hin- und Rücktransformation zwischen zeitlich aktualisierter Information (z.B. Sprache, Musik, Bewegungen) in statisch-räumliche kommt (nämlich Erinnerungsspuren, die im neuronalen Substrat fixiert sind). Lashley deutet vorsichtig die Wichtigkeit von neuronalen Rhythmen und Oszillationen an und vermutet, daß Information im Gehirn in einem Netzwerk von "reverberatory circuits" gespeichert sei. Dieser Gedanke kann in verschiedenen Versionen zum geistigen Allgemeinut gezählt werden (z.B. *Cell Assemblies* bei Hebb 1949 [zitiert nach Braitenberg 1978], kolumnäre Module bei Popper & Eccles *Phil, Bio* 1977, B-Systeme bei Braitenberg *Bio* 1978).

In (unscharfer) systemtheoretischer Terminologie könnte man sagen, daß sich bei der Informationsspeicherung schwingfähige neuronale Systeme aufbauen, die beim Informationsabruf durch Resonanzmechanismen angeregt werden. Hervorzuheben hieran ist nicht der genaue Mechanismus (in der Literatur ist eine Vielzahl davon zu finden, vgl. z.B. Grossberg & Somers *NN, Bio* 1991, Smolensky *NN, Kog* 1986), sondern das allgemeine Prinzip. Es wird nicht etwa aus einem als Gefäß vorzustellenden Speicher passiv abgelegte Information ausgelesen und an anderen Orten verarbeitet, sondern das Gedächtnis selbst ist auch der Ort der "Verarbeitung"; gespeicherte Informationen werden bei Bedarf zu Prozessen. (Auswirkungen auf das *symbol grounding* Problem diskutiert Chalmers *Phil, Kog* 1990).

Seit einigen Jahren wird in diesem Sinne im Konnektionismus die Funktion *chaotischer* neuronaler Prozesse bei der Informationsspeicherung untersucht. Yao & Freeman (*Bio, NN* 1990) rekonstruieren anhand physiologischer Befunde das Riechhirn als dynamisches System. Das Hauptresultat besagt, daß das Riechhirn im Grundzustand sich in einem hoch chaotischen Attraktor befindet und bei Reizvorgabe in schärfer definierte "Flügel" dieses Attraktors stabilisiert. Diese Flügel sind also dynamische Repräsentationen sensorischer Reize. Ähnlich wird die Bedeutung eines chaotischen Grundzustandes auch von Babloyantz & Lourenço (*Bio, NN* 1994) herausgestellt. Sie weisen darauf hin, daß ein chaotischer Attraktor im Prinzip unendlich viele semistabile Teilmuster enthält, also eine im Prinzip unbegrenzte Speicherkapazität besitzt, und sie interpretieren den chaotischen Grundzustand im Sinne eines Aufmerksamkeitszustandes (ähnlich auch Hayashi *NN, Rob* 1994).

Modelliert man kognitive Prozesse über längere Zeiträume, so muß man wiederholte Aktivierungen und Deaktivierungen verschiedener dynamischer Repräsentationen beschreiben. Tatsächlich sind in der Neuronale-Netze-Forschung und der Neurowissenschaft eine ganze Reihe von Mechanismen vorgeschlagen worden, wie neuronale Netze **Assoziationssequenzen** speichern bzw. generieren können. Ich führe einige dieser Ansätze an, um die Reichhaltigkeit der erschlossenen Phänomene und die Differenziertheit der verwendeten Techniken zu illustrieren.

Buhmann & Schulten (*NN, Phys, Inf* 1987) erklären Zeitfolgen von semistabilen Zuständen in asymmetrischen Hopfield-Netzwerken durch die Auswirkungen von Noise. Riedel et al. (*Phys* 1988) entwickeln zum selben Zweck eine analytische Theorie von Hopfield-Netzwerken mit zeitverzögernden Synapsen. Im Kontext einer komplexen, biologisch motivierten Netzwerkkonstruktion zur Musterklassifikation bauen Carpenter & Grossberg (*NN, Bio* 1990) zeitliche Verzögerungs- und Refraktäreffekte durch eine weitgehende Modellierung der chemischen Eigenschaften biologischer Synapsen ein. Reiss & Taylor (*NN, 1991*) nutzen lokale zeitliche Eigenschaften sog. *leaky-integrator*-Neurone in einem an Befunden aus dem Hippocampus orientierten Netz zur Speicherung von Zeitsequenzen. Ihre Arbeit enthält auch eine knappe Übersicht über andere Ansätze. Hayashi (*NN, Rob* 1994) benutzt ein relativ einfaches, aber chaotisch oszillatorisches Netzwerk, um Zeitsequenzen kontinuierlich ineinander überführter Muster gleichsam als "im Chaos driftende periodische Attraktoren" [intuitive Umschreibung von mir] zu lernen und zu replizieren. Yamauchi & Beer (*Rob* 1994) entwickeln mit genetischen Algorithmen rekurrente, kontinuierliche Netzwerke, welche einfache Mustersequenzen erkennen oder sich durch Reinforcement an verschiedene solche Sequenzen adaptieren können.

Die bisher in diesem Abschnitt angeführten Arbeiten betrafen überwiegend neuronale Prozesse. Der Grund für dieses Übergewicht liegt wohl einfach darin, daß Neuronale-Netz-Forscher und Biokybernetiker mit systemtheoretischen Methoden vertraut sind. Ich möchte als Kontrast mit einigen Beispielen systemtheoretischer Analysen von makroskopischen **motorischen Bewegungsmustern** aufwarten. Arbeiten dieser Art haben in der Psychophysik und der Psychologie in jüngerer Zeit systemtheoretischen Techniken zu einem ersten Durchbruch gebracht, jedenfalls in den USA (Sammelband Smith & Thelen 1993a).

Die einflußreichen Arbeiten von Schöner et al. (*Phys Psych* 1986) habe ich bereits erwähnt. Clark et al. (*Psych* 1993) beschreiben Laufbewegungen durch zyklischen Attraktoren. Durch verschiedene experimentell induzierte Störungen weisen sie die strukturelle Stabilität des Bewegungsmusters nach und beschreiben Ober- und Unterschenkel als System gekoppelter Oszillatoren. Bertenthal & Pinto (*Psych* 1993) arbeiten die These aus, wonach Menschen die Bewegungen anderer Menschen durch organisierende Prinzipien wie periodische Attraktoren, Phasenschluß u.a. wahrnehmen. Durch Experimente untermauern sie die Vermutung, daß die Ordnungsparameter, welche bei der visuellen Wahrnehmung von Bewegung entscheidend sind, dieselben sind wie die für die Generierung derselben Bewegungen wirksamen. Dies hilft, Bewegungslernen durch Imitation zu erklären. Robertson et al. (*Psych* 1993) beschäftigen sich mit den spontanen Bewegungen des Neugeborenen. Aus empirischen Bewegungstraces rekonstruieren sie einen chaotischen Attraktor in einen niedrigdimensionalen Zustandsraum. Wolff (1993) interpretiert verschiedene Wach- und Schlafzustände als Attraktoren in der Interaktionsdynamik vieler, nicht näher spezifizierter physiologischer und motorischer Subsysteme. Er belegt diese Auffassung u.a. durch detaillierte Beschreibungen von Übergängen zwischen Zuständen, die für Bifurkationen typische Merkmale aufweisen. Die Ausbildung einer Kopplung von emotionalen Zuständen und deren motorischem Ausdruck, die interkulturell wenig variiert, führt er nicht auf eine spezielle genetische Steuerung zurück, sondern auf allgemeine Prinzipien der Selbstorganisation.

Biologische Bewegungsapparate haben -- je nach Auflösungsgrad der Betrachtung -- Dutzende bis Tausende Freiheitsgrade, d.h. physikalisch nur lose gekoppelte, die Bewegung bestimmende Größen. Mit klassischen kontrolltheoretischen Methoden läßt sich ein solches System weder theoretisch modellieren noch praktisch regeln, und so war es lange rätselhaft, wie dies biologische Systeme schaffen (Smith & Thelen *Psych* 1993b). Hier bietet die Synergetik mit dem Prinzip der Versklavung eine Erklärungsmöglichkeit an. Die vielen Freiheitsgrade organisieren sich unter (mathematisch einigermaßen geklärten) Bedingungen so, daß das ganze System auf der Ebene weniger Ordnungsparameter beschrieben bzw. durch deren Variation reguliert werden kann. Dies ist ein häufig angeführtes Beispiel dafür, daß systemtheoretische Methoden dynamische Phänomene erklären, die vordem rätselhaft waren. Ähnlich kann das Versklavungsprinzip übrigens auch die koordinierte Aktivität neuronaler Populationen verständlich machen.

Ich möchte diesen Abschnitt mit einem Thema schließen, das die Rolle zeitlicher Muster bei der neuronalen Informationsverarbeitung noch einmal besonders betont. Dem Konnektionismus wurde einst vorgeworfen (Fodor & Pylyshin *Phil Kog* 1988), mit dem Phänomen der Kompositionalität kognitiver Verarbeitungseinheiten (*variable binding, feature linking*) nicht fertigzuwerden. Dieser Angriff induzierte in kurzer Zeit die Entdeckung einer ganzen Reihe von konnektionistischen Mechanismen zur Kompositionalität (vgl. van Gelder *Kog* 1990, Chalmers *Phil Kog* 1992, van Gelder & Port *Kog* 1994a). Ein heute bevorzugt untersuchter Mechanismus besteht nun gerade in global gekoppelten neuronalen Oszillationen, d.h. **Synchronisierungen** zwischen mehr oder weniger voneinander entfernten Neuronen.

Solche Synchronisationen wurden experimentell in biologischen Netzen nachgewiesen, wo sie mit der Verknüpfung einzelner sensorischer Features in Verbindung gebracht werden konnten (Engel et al. *Bio* 1990, Eckhorn & Stoeker *Bio* 1994). Stoeker & Eckhorn (*Bio* 1994) modellieren die empirischen Befunde in einem relativ einfachen neuronalen Netz in der biokybernetischen Tradition. Grossberg & Somers (*NN, Bio, Rob* 1991) konstruieren für einen erweiterten Aufgabenbereich ein komplexes "biologisch plausibles" Modell eines mustererkennenden Netzwerkes, wo ebenfalls synchronisierte Oszillationen für die Kopplung einzelner Reizkomponenten sorgen. Ein Vorzug des Synchronisierungsprinzips besteht darin, daß die Kopplung von sensorischen Features kein isoliertes Phänomen ist, sondern in einen

"ganzheitlichen" Wahrnehmungsapparat eingebunden ist und z.B. durch Aufmerksamkeitsprozesse moduliert werden kann.

Als letzter Vertreter des Prinzips Kompositionalität durch Synchronisation seien Mani & Shastri (*Kog, KI* 1993) genannt. Die Autoren präsentieren eine komplexe Netzarchitektur für Aufgaben des regelbasierten hierarchischen Klassifizierens. Variablenbindungen werden durch Synchronisation von *spike trains* geleistet. Im Gegensatz zu wohl allen bekannten klassischen KI-Systemen für denselben Zweck besitzt ihr Netzwerk im Prinzip (bei paralleler Implementierung) eine lineare Zeitkomplexität in der Länge der Eingabe, und ist beliebig erweiterbar ohne Beeinflussung der Verarbeitungsgeschwindigkeit (!). Diese Arbeit scheint mir schlagend den Nutzen einer dynamischen, d.h. prozeßorientierten Sicht auf scheinbar strukturelle Phänomene (hier: Kompositionalität) des Symbolverarbeitungsparadigmas zu demonstrieren.

### 3.2.2 Entwicklung und Lernen

Menschen zeigen in Lernprozessen oft auffällige Diskontinuitäten. Plötzliche qualitative Sprünge wechseln sich mit längeren Perioden scheinbarer Stagnation oder langsamer Reifung ab. Dieser Befund gilt sehr allgemein und auf verschiedenen Zeitskalen, unter anderem für den Spracherwerb und das Erlernen von Bewegungsmustern. Zur Erklärung wurden komplexe, durch genetische Programme zur Reifung gelangende Mechanismen für jede neu auftauchende Einzelfähigkeit oder Fertigkeitsstufe postuliert (Smith & Thelen *Psych* 1993c, Goldfield *Psych* 1993). Aus systemtheoretischer Sicht bieten sich universalere Erklärungsmöglichkeiten an, die keine speziellen Mechanismen stipulieren müssen. Drei solcher universaler Prinzipien sind Bifurkationen, nichtlineare Interaktionen von Teilsystemen, und evolutionäre Ausleseprozesse.

Bifurkationen sind ihrer Natur nach qualitative Sprünge, und ihre Allgegenwärtigkeit in dynamischen Systemen könnte viele entwicklungspsychologische Einzelbefunde durch ein einheitliches Prinzip fast zu Selbstverständlichkeiten werden lassen. Ein hübsches Beispiel für einen Entwicklungssprung durch eine Bifurkation gibt Thelen (*Psych* 1984, referiert nach Goldfield *Psych* 1993) an. Ihr gelang es, das bei Neugeborenen zu beobachtende Verschwinden einer Trittbewegung darauf zurückzuführen, daß das Bein durch Fettansatz schwerer wird, und die ursprüngliche Trittbewegung aus rein "oszillationsmechanischen" Gründen kein stabiles Bewegungsmuster mehr ist. Ein verantwortlicher Kontrollparameter konnte hier also explizit gefunden werden: das Gewicht des Beines. Goldfield (*Psych* 1993) analysiert gründlich das Krabbelverhalten bei Kleinkindern. Es gibt viele Formen des Krabbelns, die teils nacheinander, teils gleichzeitig auftreten, bei verschiedenen Kindern verschieden. Das Gesamtbild, das Goldfield zeichnet, besagt, daß die einzelnen Krabbelformen, angestoßen durch partikuläre Umwelterfordernisse, aus den gerade zur Verfügung stehenden neuralen und skeletomuskulären Elementen des Bewegungsapparates zusammengestellt werden. Wenn sich einzelne Elemente verändern, kann dies zu Bifurkationen, d.h. Reorganisationen des Bewegungsmusters, führen. Diese Sicht betont die Bedeutung von *Umweltaffordances*, Fluktuationen und Attraktoren, die eingenommen werden, ohne daß spezifische Mechanismen involviert sind. In ähnlicher Weise schlagen Tucker & Hirsh-Pasek (1993) in einem Grundsatzartikel vor, die Sprachentwicklung beim Kind zu einem großen Teil durch Bifurkationen zu erklären.

Ein anderer Erklärungsansatz für Diskontinuitäten in der Entwicklung besteht darin, verschiedene Teilfertigkeiten als nichtlinear miteinander gekoppelte und um Ressourcen (wie etwa Aufmerksamkeit) konkurrierende Teilsysteme zu interpretieren. Dieser Weg wird von van Geert (*Psych* 1993) in detaillierten Rekonstruktionen empirischer Zeitserien besonders aus dem Gebiet des Spracherwerbs besprochen. Die klassischen systemtheoretischen Gleichungen populationsdynamischer Prozesse liefern, je nach Kontrollparametern und verschiedenen anderen Details, eine Vielzahl von Wachstumskurven, die Phänomen wie den plötzlichen Durchbruch

einer Fertigkeit, Oszillationen, kurzzeitige Abschwächung einer Fertigkeit vor einem steilen Anstieg u.a.m. erfassen.

Ein dritte Sorte systemtheoretischer Modelle für Entwicklungsprozesse sind evolutionäre Mutations-Selektions-Prozesse. Hier gibt es ein weites Feld von eher praktischen Anwendungen genetischer Algorithmen zur Optimierung von technischen Systemen, unter anderem von neuronalen Netzen (vgl. Goldberg 1989, Rechenberg 1973). Im gegenwärtigen Zusammenhang sind aber vor allem gewisse Arbeiten aus der behavior-orientierten Robotik interessant (z.B. Harvey et al. AA, *Rob* 1994, Steels *KI*, AA 1994b). Hier wird versucht, die Adaptation eines *einzelnen*, konkreten Agenten an nicht a priori modellierbare Umwelanforderungen durch eine Agenten-interne Evolution zu leisten. Dabei werden die vielen Prozesse, Konzepte oder Fähigkeiten innerhalb eines Agenten als "Arten", und der Agent selbst gleichsam als "Ökosystem" aufgefaßt. Die Erfolge sind bislang nicht überzeugend. Meiner Meinung nach führt dieser Ansatz möglicherweise ins Abseits, da Agenten so etwas wie Identität besitzen sollten, Ökosysteme aber grundsätzlich "fließend" sind und sich im Laufe evolutionärer Zeitspannen beliebig verändern, sowie oft ohne klare Grenzen räumlich von einem in ein anderes übergehen.

Qualitative Entwicklungen sind -- oft unter dem schillernden Stichwort der "Selbstorganisation" (Dress et al. 1986) -- ein Hauptthema der Systemtheorie. In naturwissenschaftlich-mathematischen Rahmen sind hier inzwischen viele z.T. sehr verschiedenartige Mechanismen erforscht worden. Ihre Anwendung auf Lernen und Entwicklung von intelligenten Agenten hat jedoch, wie die Knappheit dieses Abschnitts auch zeigt, noch nicht richtig begonnen. Für die Zukunft sind hier meiner Meinung nach jedoch außergewöhnlich stimulierende Beiträge zu erwarten.

### 3.2.3 Konzepte

In der Kognitionswissenschaft, der KI und der Linguistik werden Konzepte meistens extensional als Klassen von Referenten interpretiert (Smith & Medin *Ling, Kog* 1981). Diese Sicht bereitet jedoch Schwierigkeiten mit einer ganzen Reihe von empirischen Phänomenen (Treisman *Psych* 1986, Medin *Psych* 1989, Goschke & Koppelberg *Psych, NN* 1990). In diesem Abschnitt sollen einige Schwierigkeiten der extensionalen Sichtweise angesprochen werden, die sich mit systemtheoretischen Interpretationen auflösen lassen. Dazu wird zunächst referiert, was es zur Zeit an systemtheoretisch orientierten Konzeptmodellierungen gibt.

Die wohl größte Klasse dieser alternativen Modellierungen benutzt lokalistische neuronale Netzwerke, in denen Knoten mit Konzepten, Features, syntaktischen Merkmalen, aber auch übergeordneten Kontexten benannt sind. Gleichgewichtszustände (d.h. Fixpunktattraktoren) von Aktivierungen in solchen Netzwerken werden als Konzeptaktualisierungen interpretiert. Die Tradition solcher Netzwerke läßt sich bis auf Quillians (*KI*, 1968) semantisches "Ur-" Netzwerk und Wilks' (*KI, Ling* 1975) *spreading-activation*-Netzwerk zurückverfolgen. Der konnektionistische Gedanke einer sich äquilibrierenden Aktivierung kam jedoch erst bei Waltz & Pollack (*Kog, Ling* 1985) voll zum Zuge. Seine Arbeit hat bei einer ganzen Reihe von Nachfolgern Verfeinerungen erfahren. Hier sei Mangold-Allwinn (*Ling* 1991) genannt, dessen Arbeit einen ausgezeichneten Überblick enthält. Smolenskys (*NN* 1986) *harmony theory* beruht ebenfalls auf Aktivierungsäquilibration in einem lokalistischen Netzwerk. Im Unterschied zu den vorgenannten Arbeiten leistet der Autor jedoch eine gründliche mathematische Analyse mit Mitteln der Systemtheorie. Legendre et al. (*Ling, NN* 1990a,b) erweitern diesen Ansatz von einzelnen Konzepten auf komplexere linguistische Strukturen.

Systemtheoretisch gesprochen, werden in diesen Arbeiten Konzepte als Fixpunktattraktoren modelliert. Die Trajekturen dieser Systeme konvergieren in der Regel rasch zu einem Punktattraktor. Dem entspricht die Verwendung der Netze für Klassifikationen eines einzelnen vorgegebenen Musters. Ein *Inputstrom* ist nicht vorgesehen.

Auf einige Arbeiten, die bereits zitiert wurden, muß an dieser Stelle noch einmal hingewiesen werden. Auch dort wurden (sensomotorische) Konzepte als Punktattraktoren modelliert, zumeist in neuronalen Netzen: bei Braitenberg (1978), Riedel et al. (1988), Buhmann & Schulten (1987), Carpenter & Grossberg (1990), und Steels (1994a).

Grenzykel und chaotische Attraktoren sind ebenfalls als Modell für Konzepte vorgeschlagen worden. Aus vorangegangenen Abschnitten sei hier noch einmal an Babloyantz' & Lourenços (1994) semistabile Grenzykel, sowie Yao & Freemans (1990) chaotische Attraktoren erinnert.

In Jaeger (*KI, Kog* 1991, 1992, 1994a,d) entwickle ich dynamische Konzeptrepräsentationen, die auf selbststabilisierenden, zyklischen Graphstrukturen beruhen. Sie lassen sich als komplexe Attraktoren in einem diskreten, räumlich verteilten System deuten. Das Modell (*dynamische Symbolsysteme* getauft) erfaßt simultane top-down und bottom-up Interaktionen zwischen Granularitätsstufen, die von der sensomotorischen Peripherie bis zu einem Konzepte verarbeitenden Zentrum reichen, und erklärt Abstraktionsbeziehungen durch eine Art Bifurkation.

Wenn Konzepte durch Attraktoren modelliert werden, kann man die zugehörigen Attraktionsbecken als eine Art Extension interpretieren. Ein Attraktionsbecken zu einem Attraktor A enthält genau die Systemzustände, aus denen sich das System zum Attraktor A hinentwickelt. Man kann also (Start-)Zustände eines Systems danach *klassifizieren*, auf welchen Attraktor sie sich hinentwickeln. Entsprechen die Startzustände dann vorgegebenen (z.B. sensorischen) Inputinformationen, so werden auf diese Weise Inputmuster klassifiziert. Balkenius und Gärdenfors (1991) analysieren in diesem Sinne die Klassifikationsleistungen bestimmter neuronaler Netze vor einem klassischen logischen Hintergrund, und Happel & Murre (*NN* 1995) machen die interessante Beobachtung, daß sich bei unüberwacht gelernten Klassifikationen in neuronalen Netzen fraktal begrenzte Attraktionsbecken, d.h. sehr komplexe, vergrößert jedoch als "fuzzy" erscheinende Extensionen, bilden können.

Neben Attraktoren findet man in der Literatur noch eine völlig andere Art, Konzepte (oder allgemeiner, diskrete Verarbeitungseinheiten) in dynamischen Systemen zu realisieren. Man verwendet dazu einfach bestimmte abgeschlossene Teilvolumen des Phasenraumes, sog. *Partitionszellen*. Ein solchermaßen repräsentiertes Konzept wird aktualisiert, wenn und solange die Systemtrajektorie durch die entsprechende Partitionszelle verläuft. Diese Sicht ist charakteristisch für die informationstheoretische, von theoretischen Physikern betriebene Erforschung von Systemen, welche Symbolfolgen generieren (z.B. Crutchfield & Young *Math, Phys* 1990, J.S. Nicolis & Tsuda *Math, Phys* 1989, Ebeling & G.Nicolis *Ling, Phys* 1992). Ferner ist sie die Grundlage für Arbeiten, in denen neuronale Netze die Leistungen endlicher Automaten zu simulieren lernen bzw. neuronale Netze durch endliche Automaten rekonstruiert werden (z.B. Tani *AA, NN* 1995, Tino et al. *Math, NN* 1995, Giles und Omlin 1995). Diese Forschungen erhellen den Übergang zwischen nichtsymbolischer, kontinuierlicher und symbolischer, diskreter Informationsverarbeitung, welches ein Teilaspekt des *symbol grounding* Problems ist (Harnad *Psych, Phil* 1990, Chalmers *Kog, Phil* 1990).

An dieser Stelle sei auch angemerkt, daß in der Forschung über qualitatives Schließen diskrete Partitionierungen von kontinuierlichen Zustandsräumen die Standardprozedur ist, um diskrete Ereignisfolgen zu erhalten (z.B. Ling & Buchal *KI* 1993).

Attraktor- und Partitionszellenmodelle für Konzepte sind komplementär: erstere erklären auf natürliche Weise die *Stabilität* von Konzepten, und letztere die Tatsache, daß Konzepte in *Aktivierungssequenzen* vorkommen. Attraktoren können per definitionem nicht in Sequenzen auftreten, da sie eben dadurch definiert sind, daß sie die Systemtrajektorie für  $t \rightarrow \infty$  "einfangen". Will man (eigentlich ein Widerspruch in sich) Folgen von Attraktorzuständen erhalten, muß man Zusatzmechanismen einführen (vgl. die Angaben zu Assoziationssequenzen in 3.2.1). In Jaeger (*AA, Math* 1995c) schlage ich ein mathematisches Konstrukt vor (*transiente Attraktoren*), welches als eine informationstheoretisch gefaßte Kreuzung zwischen

Attraktoren und Partitionszellen gesehen werden kann. Transiente Attraktoren vereinigen ohne Zusatzmechanismus den Stabilitäts Gesichtspunkt mit der Möglichkeit von Aktualisierungssequenzen. Die Arbeit gibt auch ein effektives Verfahren an, wie empirisch gefundene Phasenbilder als Generatoren von Assoziationsfolgen transienter Attraktoren analysiert werden können.

Einige weitere, in Einzelarbeiten auftretende dynamische Repräsentationsformate für Konzepte werden in van Gelder (*Phil, Kog* 1995) aufgeführt.

Insgesamt ergibt sich aus dieser knappen Übersicht, daß es für Konzepte in kontinuierlichen dynamischen Systemen formale Korrelate gibt, welche dynamische Eigenschaften von Konzepten erst überhaupt zu problematisieren und ansatzweise auch schon zu erklären gestatten, und daß es allgemeiner schon eine ganze Reihe konkreter Techniken gibt, wie man kontinuierlich-quantitative als symbolverarbeitende Systeme interpretieren kann.

Im Rest dieses Abschnittes möchte ich verschiedene Merkmale empirischer Konzepte behandeln, welchen man sich mit logikorientierten Methoden schlecht und mit systemtheoretischen gut nähern kann.

Empirisch beim Menschen beobachtete Konzepte sind **zeitvariabel**. Sie verändern sich im Maßstab von Generationen (Lakoff *Ling* 1987), Jahren und Monaten (Bartlett *Psych* 1932), Wochen (Barsalou *Ling, Kog* 1987), Zehntelsekunden (Seidenberg et al. *Psych* 1982), in Bezug auf verschiedenste meßbare Aspekte. Das läßt sich nicht gut mit einer logik-basierten Konzeptmodellierung erfassen, da logische bzw. modelltheoretische Bestimmungen inhärent zeitlos sind (Diskussion in Jaeger *KI, Math* 1992, 1994a, p.134ff). Dagegen nehmen systemtheoretische Modellbildungen gerade von der Zeitlichkeit der betrachteten Systeme ihren Ausgang. Ich möchte diese Grundtatsache hier noch einmal hervorheben.

Temporale Phänomene reichen von kurzfristigen Äquilibrierungs-, Synchronisierungs- und Oszillationsvorgängen über mittelfristiges Lernen und Adaptation bis zu Generationen überdauernden, evolutionären Prozessen, wozu es verschiedene systemtheoretische Mechanismen gibt. Eine einheitliche Theorie der phänomenalen Zeitlichkeit in dynamischen Systemen ist nicht in Sicht und nach allem, was wir von dem hier waltenden Erscheinungsreichtum heute schon ahnen können, wohl auch nicht möglich. Komplexere Ansätze zur dynamischen Modellierung kognitiver Systeme (konnektionistisch bei Carpenter & Grossberg 1990 und Gaudiano & Grossberg 1991, Yao & Freeman 1991, mit evolutionären Techniken bei Drescher 1991, Holland & Reitman 1978, genial eklektisch bei Hofstaedter & Mitchell 1993, mit dynamischen Symbolsystemen bei Jaeger 1994a) enthalten mehrere zeitliche Mechanismen auf verschiedenen Skalen und demonstrieren einerseits die Beschränktheit der heute möglichen theoretischen Durchdringung, aber andererseits auch die Möglichkeit, die zeitliche Phänomenologie von Konzepten überhaupt zu modellieren.

Eine Art von **Kontextabhängigkeit** von Konzepten äußert sich darin, daß beim Aufruf eines Konzeptes in Abhängigkeit von gleichzeitig oder vorher aktualisierten Konzepten verschiedene Merkmale salient werden. Hierzu gibt es hierzu viele empirische Untersuchungen (Mangold-Allwin *Ling* 1991, Barsalou *Kog* 1987, 1989, Treisman *Psych* 1986 p. 35-44, Richards *Psych* 1988, Goschke & Koppelberg *Ling, NN* 1990). Um das Problem in einem klassischen methodischen Rahmen bei den Wurzeln zu packen, muß man die gewohnte Prädikatenlogik aufgeben und zum Beispiel Situationssemantiken oder nichtmonotone Logiken einsetzen. Dadurch handelt man sich allerdings im allgemeinen rechnerische Intraktabilität ein. In Anwendungen hilft man sich oft mit pragmatischen "Hacks" (Beispiele in Brézillon *KI* 1993).

Systemtheoretisch lassen sich Kontexteinflüsse auf mehrere Weisen einfach modellieren. Die gängigste Art ist in lokalistischen neuronalen Netzwerken zur "one-shot"-Klassifikation verwirklicht, in denen Konzepte Punktattraktoren sind (s.o.). Hier tauchen Kontexteinflüsse als

Kontrollparameter auf (in der Form von festgehaltenen [engl. *clamped*] Inputknoten-Aktivierungen). In einem gewissen Bereich führt die Variation dieser Kontrollparameter zu stetigen Verschiebungen der Attraktoren; bei Abweichungen über kritische Werte hinaus kommt es zu Bifurkationen. Dem entsprechen eine kontextabhängige, stetige Verschiebung von Konzeptinterpretationen und plötzliche Umdeutungsereignisse. Eine solche qualitative Differenz ist zum Beispiel die Interpretation von *buck* als Hirsch vs. Dollarnote bei Waltz & Pollack (*Kog, Ling* 1985).

Das Prinzip "Kontext durch Kontrollparameter" läßt sich natürlich auch bei komplexeren Attraktormodellen für Konzepte verwenden. Eine besonders interessante Variante könnte die von Babloyantz und Laurenço (*Bio* 1994) vorgeschlagene Konzeptmodellierung durch reizstabilisierte zyklische Attraktoren in einem chaotischen Grundzustand abgeben. Je nach Reizzusammensetzung können verschiedene Zyklen simultan stabilisiert werden. Eine Konzeptaktualisierung ist hier also eine spontane, Input-getriebene Zusammenstellung von etwas, das man Features nennen könnte. Diese ist kontextsensitiv, insofern der Input Kontextinformation enthält.

Am direktesten werden Kontexteinflüsse vielleicht in räumlich verteilten Systemen verwirklicht, in denen einzelne Konzepte *lokalisiert* aktualisiert werden und dadurch *räumlich* Kontexte füreinander bilden. In Gehirnen ist die räumliche Struktur funktional. Manchmal lassen sich Konzepte kleinräumigen Arealen zuordnen (z.B. von Frucht- und Gemüseamen, berichtet von Hart et al. *Psych, Bio* 1985). Im visuellen Kortex gibt es eine kleinräumige Feature-Spezialisierung von Neuronenverbänden, deren Genese mit *self-organizing feature maps* in künstlichen Netzwerken rekonstruiert werden kann (Obermeyer et al. *NN, Bio* 1990). Neben solchen Befunden einer räumlichen Lokalisierung von Konzepten oder verwandten Phänomenen gibt es jedoch bekanntermaßen auch Befunde, die auf eine großräumig verteilte, "holistische" Repräsentation deuten. Insgesamt kann man wohl davon ausgehen, daß es in Gehirnen mehrere räumliche Ordnungsprinzipien gibt, die auf noch kaum geklärte Weise nach Hirnareal, Konzeptsorte u.a. variieren. Außerdem sind Wechselwirkungen räumlicher Organisation mit zeitlichen Phänomenen wichtig, aber noch nicht systematisch untersucht. Erste Ansätze sind z.B. die räumliche Speicherung zeitlicher Sequenzen in einem Kohonen-Netz mittels *leaky-integrator*-Neuronen (Campbell & Taylor 1993) und das schon in 3.2.1 erwähnte Netzwerk von Matsuga & Yuille (*NN*, 1994).

Eine explizite Kontextmodellierung in einem räumlich verteilten System war auch eines der Grundmotive für meine Entwicklung der dynamischen Symbolsysteme (Jaeger 1992, 1994a).

**Zyklische, iterierte und selbstreferentielle Beziehungen** zwischen Konzepten sind verbreitet und faszinierend. Das Buch, das wohl mehr als jedes andere junge Studenten zur KI und Kognitionswissenschaft gebracht hat, Douglas Hofstadters "Gödel, Escher, Bach" (*KI, Kog* 1979), ist eine Orgie der Selbstbezüglichkeit. Versucht man, konzeptuelle Informationsverarbeitung in logikorientierten Formalismen wie etwa KL-ONE-artigen Sprachen zu implementieren, so bringt einen diese Selbstbezüglichkeit rasch in Schwierigkeiten. Sie gründen praktisch in algorithmischen Problemen mit Zykeln, und theoretisch darin, daß man für die modelltheoretische Semantik eine fundierte Mengenlehre verwendet (Nebel *KI* 1990, 1991, weitere Diskussion in Jaeger *KI, Math* 1994a, aber: nichtfundierte Mengenlehre zur Modellierung von *mutual beliefs* bei Barwise *Math, Kog* 1989).

Demgegenüber sind Selbstbezüglichkeit und Zyklizität in Formalismen dynamischer Systeme von vorneherein eingebaut. Hier muß ich mich nun doch einmal auf Formeln beziehen. In Differentialgleichungen (d.h. Formeln der Art  $dx/dt = f(x)$ ) oder diskreten Abbildungen (Formeln der Art  $x_{t+1} = f(x_t)$ ), den beiden gebräuchlichsten Hilfsmitteln zur Spezifikation dynamischer Systeme, wird die Veränderung der Systemgrößen ( $dx/dt$  bzw.  $x_{t+1}$ ) auf einen (im Falle von Differentialgleichungen infinitesimal) früheren Wert *derselben* Systemgrößen

zurückgeführt. (Schein-)Probleme der Zyklizität und der Selbstbezüglichkeit werden so in systemtheoretischer Sicht durch eine Würdigung der Rolle der Zeit aufgelöst bzw. treten erst gar nicht auf (von Foerster 1984). Dynamische Systeme entwickeln sich aus sich selbst weiter - - das ist gar nicht enigmatisch, sondern eine knappe Charakterisierung der Wirkungsweise von Differentialgleichungen etc. (eingehendere Diskussion in Jaeger 1994a).

In der kognitionswissenschaftlichen, linguistischen und der KI-Literatur wird Konzepten gerne *en passant* **Gestalteigenschaften** zugesprochen. Nun ist der Gestaltbegriff nicht präzise operationalisiert. In der klassischen Gestaltpsychologie (Nyman *Psych* 1966) sind zwar viele empirische Untersuchungen zur Wahrnehmung von Gestalten gemacht worden. Zentrale Begriffe wie *Abgesetztheit*, *Geschlossenheit*, *Gestalthöhe/-reichtum*, *Gestaltgüte/-festigkeit*, *Gerichtetheit*, *Gegliedertheit*, *Stabilität* und die sog. *Gestaltgesetze* sind jedoch weder formal, noch unabhängig von bestimmten Aufgabenkontexten, noch unabhängig von Introspektionsmethoden definiert worden. So ist es kein Wunder, daß "Gestalthaftigkeit" immer noch eher eine Metapher als eine Erklärung ist.

Es gibt viele lockere Querverbindungen zwischen Systemtheorie und Gestaltpsychologie. Von Bertalanffy (*Bio* 1968 p.208) sieht seine allgemeine Systemtheorie historisch als Parallelentwicklung zur deutschen Gestaltpsychologie. Wiener (*Ing, Math, Bio, Kog* 1961) widmet der visuellen Gestalterkennung ein eigenes Kapitel. Wiener arbeitete zeitweise mit dem Gestaltpsychologen Kurt Lewin zusammen. Eigen & Winkler (*Bio* 1975, Kap. 6) stellen in einem lockeren Übersichtskapitel über verschiedene chemische und biologische Strukturbildungen wiederholt Querverbindungen zum Gestaltbegriff her. Umgekehrt wird aus der Psychologie heraus gelegentlich eine Verbindung von der Gestaltpsychologie zu systemtheoretischen Gedanken hergestellt (in Nebenbemerkungen bei Krause *Psych* 1989, zentral und ausführlich bei Stadler & Kruse *Psych* 1986, 1990).

In allen aufgeführten Arbeiten bleiben die Querbezüge zwischen Gestalttheorie und naturwissenschaftlicher Systemtheorie oberflächlich. Hier wird meiner Meinung nach eine doppelte Chance vertan. Erstens wäre die in der Gestaltpsychologie erschlossene, reiche Phänomenologie von Perzepten und Konzepten für ein systemtheoretisches Konzeptmodell nutzbar zu machen. Die in systemtheoretischen (zumeist konnektionistischen) Ansätzen der Kognitionswissenschaft und der KI verbreitetete, aber wohl zu einfache Modellierung von Konzepten durch Attraktoren (vgl. Abschnitt 3.3) könnte durch die Aufnahme des in der Gestaltpsychologie gesammelten Materials differenziert und angereichert werden. Umgekehrt könnten systemtheoretische formale Modelle zu der sehr wünschenswerten Präzisierung des Gestaltbegriffs in der Psychologie führen. Solche Modelle könnten insbesondere den Prozeßcharakter von Gestaltwahrnehmungen erschließen. Einen allerersten Schritt zu einer Formalisierung und Dynamisierung gestalttheoretischer Konzepte unternehme ich in Jaeger (*Math, KI, Kog* 1994b).

### 3.2.4 Agent-Umwelt-Systeme

Die Natur vieler Intelligenzleistungen kann nur verstanden werden, wenn man die physikalische Körperlichkeit von intelligenten Agenten und die physikalische Interaktion mit der realen Umgebung berücksichtigt. Diese Einsicht ist in der traditionell einer entkörperlichten Intelligenzauffassung verbundenen Kognitionswissenschaft recht jung und keineswegs ausdiskutiert. In der *situated action* Debatte in der KI und der Kognitionswissenschaft (z.B. Suchman *Kog, KI* 1987, Vera & Simon *KI* 1993a, Clancey *KI* 1993) werden verschiedene Aspekte dieser Perspektive teilweise in ungewöhnlich polemischer Weise (z.B. Vera & Simon 1993b, Hayes et al. *KI, Phil* 1994) diskutiert. Für die behavior-orientierte Robotik ist die Körperlichkeit und phy-

sikalische Situiertheit von Robotern der zentrale methodologische Ausgangspunkt (Brooks *KI, AA, Rob* 1991, Steels *KI, AA* 1993a).

Wenn man formale Modelle körperlicher und physikalisch situierter Agenten aufstellen will, bieten sich systemtheoretische Techniken kanonisch an.

Was den Aspekt der Körperlichkeit betrifft, so *sind* Roboter- und Tierkörper mechanische, dynamische Systeme. Dementsprechend sind systemtheoretische (genauer, kontrolltheoretische) Methoden in der Robotik Standard, und die Psychophysik ist die Sparte der Psychologie, in der solche Methoden am ehesten eingesetzt wurden (vgl. die Bemerkungen zur Bewegungsmodellierung in 3.2.1). Akzentuiert wird die Verbindung der Körperlichkeit mit systemtheoretischen Methoden in einem Positionspapier von Dautenhahn & Christaller (*AA, KI* 1995).

Ich möchte hier mehr auf den zweiten Aspekt eingehen, den der physikalischen Situiertheit. Obwohl physikalische Situiertheit ein Leitmotiv für die Paradigmen der *situated action* und der behavior-orientierten Robotik ist, so ist noch nicht ganz klar, was dieser Terminus eigentlich bedeuten soll und welche Konsequenzen die damit verbundenen Ideen für die Erforschung der Intelligenz haben. Hier sollen einige Aspekte dieses Komplexes aus systemtheoretischer Sicht näher beleuchtet werden.

Eine besonders konsequente Version des Situiertheitsgedankens besteht darin, Agent und Umwelt als ein einziges dynamisches System aufzufassen. Dies wird von manchen Vertretern der behavior-orientierten Robotik getan (Smithers *AA* 1994, Beer *KI, AA, NN* 1995). Hier wird vorgeschlagen, Agent-Umwelt-Systeme durch Differentialgleichungen zu beschreiben, wobei der Agent und die Umwelt jeweils formal als ein Teilsystem aufgefaßt werden.

Diese Arbeiten sind Prinzipienpapiere. Es scheint mir jedoch aus vielerlei Gründen ausgeschlossen, mit Differentialgleichungen eine nichttriviale Umwelt zu modellieren. So stellt sich die Frage nach vereinfachten, diskret approximierenden, evtl. stochastischen Systemmodellen insbesondere für die Modellierung der Umwelt in einem Agent-Umwelt-System. In diese Richtung gehen Kaelbling und Mitarbeiter (Entwurf und Motivation in Kaelbling *KI* 1991, Ausführung in Baayse et al. *KI* 1995). Sie fassen die Umwelt als endlichen Automaten auf, in dem Aktionen eines Roboters für Zustandsübergänge sorgen. Für einen Roboter stellt sich die Aufgabe, unter Bedingungen sensorischer und aktuatorischer Unsicherheit aus Aktionssequenzen ein Modell der Umwelt zu lernen. Die Autoren geben hierfür praktikable Algorithmen an. Es wäre eine spannende Aufgabe, solche diskreten Umweltmodelle mit formal gleichartigen Agentenmodellen (z.B. Košecká & Bajcsy *Rob* 1994) zu umfassenden diskreten Modellen von Agenten-Umwelt-Systemen zu integrieren.

Nimmt die Kognitionswissenschaft den Situiertheitsgedanken ernst, so wird sie eindringlich mit Fragen der Bewegungskontrolle und deren sensorischer Steuerung konfrontiert. In diesem bislang eher vernachlässigten Bereich kann sie viel von der Robotik profitieren, wo die Regelung von Effektorbewegungen und räumlicher Fortbewegung eine Grundaufgabe darstellt.

Ich will hier nur ein Beispiel aus der umfangreichen Literatur angeben. Triggs (*Rob* 1994) beschreibt die Navigationssteuerung eines mobilen Roboters auf der Grundlage von Sonarsensoren. Diese haben eine geringe Winkelauflösung und sind anfällig für Fehlinterpretationen durch Mehrfachreflexionen. Diese Schwierigkeit wird weitgehend durch klassische kontrolltheoretische Techniken bewältigt, welche eine detaillierte formale Modellierung der geometrischen, physikalischen und stochastischen Eigenschaften der verwendeten Sensoren und der häufigsten Umweltelemente (!) einschließt.

Die Lektüre solcher Arbeiten ist auch für Kognitionswissenschaftler sehr lohnend, da sensomotorische Komponenten von Objektbegriffen äußerst klar hervortreten. Dieselbe Grundhaltung findet sich in verschiedenen Ausprägungen auch in einer Reihe von Arbeiten aus verschiedenen anderen Gebieten wieder, wo mit *interaktionistischen* Konzeptbegriffen gearbeitet wird. Der repräsentationale Gehalt von Konzepten wird hier nicht (wie in der Modelltheorie) in einer ideellen Referenzbeziehung zwischen Begriff (bzw. dessen Symbol) und externem Denotat gesehen. Vielmehr resultiert der repräsentationale Gehalt eines

Konzepts aus Invarianten in der Interaktionsgeschichte eines Agenten mit äußeren Objekten. "Konzepte" und "repräsentierte Objekte" sind wechselseitig aufeinander angewiesen; beide zusammen *sind* ein einziges dynamisches Interaktionsmuster. Versionen dieser Auffassung sind in der behavior-orientierten Robotik verbreitet (Verschure AA, *Rob* 1993, Scheier & Pfeifer AA, *Rob* 1995), aber auch in der Biokybernetik und konnektionistischen Robotik (Gaudio & Grossberg *Bio, Rob, NN* 1991 und Zalama et al. *Bio, Rob, NN* 1995), in der Theorie autonomer Agenten und in der Epistemologie der Kognitionswissenschaft (sehr profund: Bickhart *Phil* 1993), in der KI (Drescher *KI* 1991), und natürlich in der Entwicklungspsychologie Piaget'scher Prägung (Flavell *Psych* 1968).

Von besonders eifrigen Vertretern des *situated-action*-Paradigmas wird die direkte, sensomotorische Rückkopplung vom Agenten in seine Umgebung in einer Weise interpretiert, welche die Existenz von symbolischen Repräsentationen überhaupt infrage stellt. Ich möchte hier ausdrücklich feststellen, daß eine systemtheoretisch inspirierte Sicht auf situierte Agenten nicht nur gut mit der Existenz von internen Repräsentationen verträglich ist, sondern sogar zu einer besonders reichen und interessanten, nämlich dynamischen Auffassung von Repräsentationen führen kann (Diskussion in van Gelder *Phil, Kog* 1995).

### 3.2.5 Semantik

Die logische Modelltheorie und die klassische Linguistik spielen sich in einem platonischen Raum ab, wo es extensional bestimmte Klassen von Gegenständen einfach "gibt", und wo Symbole diese Klassen einfach "bedeuten" können. Dabei bleibt der intensionale Aspekt von Konzepten, wie sie konkret von einem individuellen Menschen verwendet werden, rätselhaft. Die Situationssemantik (Devlin *Math, Ling* 1991) stellt sich zwar der Problematik individueller und situativer Konzeptualisierungen, jedoch stipuliert auch sie ohne weitere Erklärung die Existenz von Referenzbeziehungen zwischen Symbolen (in Köpfen) und Objekten (in der äußeren Realität). Auch sie versucht nicht, die physikalischen und physiologischen Mechanismen zu klären, die einen Menschen (oder ein Tier oder vielleicht einen Roboter) dazu befähigen, "aua" zu sagen, wenn er mit dem Fuß gegen den Bordstein tritt. Grob gesagt, klassische Modelltheorien und Situationssemantik sind Werkzeuge für Philosophen, nicht für empirische Wissenschaftler oder Ingenieure (schöne Darstellung und Diskussion in Woolridge & Jennings *KI, Inf* 1995).

Dank ihrer naturwissenschaftlichen Herkunft faßt eine systemtheoretische Sicht von situierten Agenten diese durch und durch als physikalische Individuen in einer physikalischen Umwelt auf. Von dem Impuls, den der Bordstein dem Fuß gibt, über Reizleitungs- und Integrationsmechanismen bis zur Ausbildung eines komplexen Attraktormusters im Neokortex mit Aktivierung des Stimmapparats verbleibt die Rekonstruktion des Ereignisses in einem einzigen ontologischen Rahmen, nämlich dem physikalischen (Diskussion in Jaeger 1994a, Kap. 2.4, 1994d). Hier ist gar kein Platz für die Idee eines Symbols, das als solches eine "Bedeutung" trüge. Zu Symbolen werden die im Neokortex der Vpn beobachteten Konzeptualisierungsergebnisse erst in der Metasprache des Forschers, wenn er sie seinerseits klassifiziert und in Publikationen durch Schriftzeichen belegt. Als *Objekte* der Theoriebildung kommen Symbole in systemtheoretischen Modellen informationsverarbeitender Systeme gar nicht vor. Kurz, ein situiertes Agent ist in systemtheoretischer Sicht nichts anderes als ein komplexes, aber durch und durch physikalisches Meßgerät: gegeben den und den physikalischen Input, liefert es den und den physikalischen Zeigerausschlag ("aua").

Diese Sicht auf das Problem der Semantik harmoniert bestens mit der Körperlichkeit und physikalischen Situietheit von Agenten, und sie erfaßt sofort die Zeitlichkeit von situierten Informationsverarbeitung. Sie ist naturgemäß recht aufwendig in der konkreten Anwendung, da der ganze sensomotorische "Meßapparat" mit berücksichtigt werden muß, der in der plato-

nischen Symbolsemantik glatt ignoriert wird. Die Modelle kompletter neuronaler Systeme, die in Jahrzehnten etwa von Freeman und Mitarbeitern (z.B. Yao & Freeman 1990) oder Grossberg und Mitarbeitern (z.B. Carpenter & Grossberg 1990) erarbeitet worden sind, zeigen aber, daß solche Komplettrekonstruktionen heute bereits möglich sind, und setzen damit Standards wohl auch für die Kognitionswissenschaft.

Die hier angedeutete systemtheoretische Semantiktheorie ist von Crutchfield (*Phys, Math* 1992) bereits formaler ausgearbeitet worden. Auf informationstheoretischer Grundlage analysiert er die semantische Beziehung, die durch einen physikalischen Meß/Interpretationsprozess zwischen einem physikalischen Prozess und einem durch den Interpretationsprozeß gelieferten einfacheren "Modell" des Prozesses besteht.

Einen gänzlich anderen Zugang zu einer "dynamischen Semantik" weist der Logiker Gärdenfors (*KI, Math* 1994). Er definiert die Bedeutung einer Proposition durch die Veränderung, welche sie im Glaubenssystem eines Agenten bewirkt, falls dieser die Botschaft akzeptiert. Propositionen werden hier als Funktionen auf *information states* definiert. Von diesem Grundgedanken aus rekonstruiert Gärdenfors die Schlußregeln einer vollständigen propositionalen Logik. Diese Sichtweise auf eine propositionale Semantik ist einerseits dynamisch (weil sie auf zeitlichen Veränderungen beruht) und intensional und somit sehr verschieden von der herkömmlichen modelltheoretischen Semantik der Logik. Andererseits ist diese Perspektive auf Satzbedeutungen auch klassisch, insofern sie sich in einem ideellen Raum von "körperlosen" Informationen bewegt.

### 3.3 Neuronale Netze und Dynamische Systeme

Der Titel dieses Aufsatzes verspricht, von dynamischen Systemen zu handeln. Ein Großteil meiner Beispiele betrifft jedoch neuronale Netze. In einem Satz ausgedrückt, besteht hier folgende Verbindung: *rekurrente neuronale Netze und dynamische Systeme mit reellwertigen Systemgrößen sind im wesentlichen aufeinander abbildbar*. Entsprechend schlagen Munro & Anderson (*NN* 1988) auch vor, die akademische Lehre im Bereich neuronale Netze durchgängig auf der Interpretation von neuronalen Netzen als dynamischen Systemen aufzubauen.

Wichtig ist hier das Attribut "rekurrent". Ein neuronales Netz heißt rekurrent, wenn es Rückkopplungen enthält, d.h. wenn die Aktivierungen gewisser Neurone (evtl. auf Umwegen) auf dasselbe Neuron zurückwirken. Erst solche Rückkopplungen befähigen das Netz zur Aufrechterhaltung einer autonomen Dynamik. Dies steht im Gegensatz zu nicht rekurrenten *feed-forward*-Netzen, die einen gegebenen Input ein Mal in einer Richtung "durchschleusen" und danach ein Ergebnis abliefern. Mathematisch gesehen verwirklichen solche Netze *Funktionen*.

Ohne hier auf Details einzugehen, möchte ich den engen Zusammenhang zwischen rekurrenten neuronalen Netzen und dynamischen Systemen durch ein Diagramm verdeutlichen. Abb. 5 zeigt eine Skizze eines hierarchischen Systems, das sowohl als neuronales Netz als auch als dynamisches System gelesen werden kann. In der ersten Lesart entsprechen die Knoten  $x$ ,  $y_i$ ,  $z_j$  den Neuronen, die Pfeile den synaptischen Verbindungen. Zu jedem Neuron wäre eine *Aktivierungsfunktion* anzugeben, die spezifiziert, wie sich die Aktivierung des Neurons in Abhängigkeit der Aktivierungen derjenigen anderen Neurone verändert, von denen es synaptischen Input erhält. Die Aktivierungsfunktion für das Neuron  $y_2$  etwa hätte die Form  $y_2(t + 1) = f(x(t), z_2(t), z_3(t), z_4(t))$ . Als Architekturskizze für ein dynamisches System gelesen, wären die  $x$ ,  $y_i$ ,  $z_j$  die Systemgrößen, und die Pfeile entsprächen den Abhängigkeiten, wie sie sich in den Systemgleichungen ausdrücken. Die Systemgleichung für die Systemgröße  $y_2$  etwa hätte wieder die Form  $y_2(t + 1) = f(x(t), z_2(t), z_3(t), z_4(t))$ .

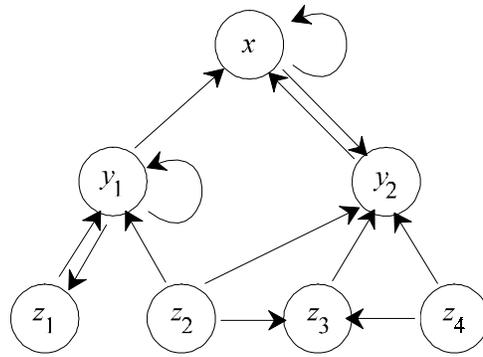


Abb. 5: Architekturskizze eines neuronalen Netzes bzw. dynamischen Systems.

Auch wenn rekurrente neuronale Netze und dynamische Systeme rein formal sehr ähnlich sind, so gibt es natürlich doch wichtige Unterschiede in der Praxis. In der Theorie neuronaler Netze beschränkt man sich auf sehr spezielle Klassen möglicher Aktivierungsfunktionen, und auf sehr wenige Topologien für Architekturen wie in Abb. 5. Außerdem beschäftigt man sich zentral mit Lernmechanismen. Diese könnten als der eigenständige, nicht direkt auf Standardkonstrukte der Systemtheorie abbildbare Anteil im Konnektionismus gelten. Eigentlich systemtheoretische Beschreibungen von Lernprozessen in neuronalen Netzen sind selten (z.B. als Bifurkationskaskade in Tani 1995).

Abschließend möchte ich noch ein Punkt herausstellen, der vielleicht einem gewissen Vorurteil widerspricht. Naturwissenschaftlich-mathematische Theorien dynamischer Systeme beschäftigen sich mit *qualitativen* Eigenschaften von Systemen, wie Attraktoren, Bifurkationen etc. Dem entsprechen in neuronalen Netzwerken "globale" Phänomene: Aktivierungsmuster als ganze, Folgen davon, chaotische globale Zustände, etc. Es geht im Konnektionismus zumeist gar nicht um die elementare "hardware"-Ebene im neurologischen Detail (dazu sind künstliche neuronale Netze auch viel zu sehr vereinfachende Abstraktionen), sondern um globalere Funktionsprinzipien. Die Abstraktionsebene in der neuronale-Netze-Forschung ist höher als vielleicht von außerhalb wahrgenommen wird.

### 3.4 Wie erklären systemtheoretische Modelle?

Naturwissenschaftliche Erklärungen sind typischerweise reduktionistisch: ein Phänomen wird durch Angabe eines "zugrundeliegenden" Mechanismus erklärt. Im Gegensatz dazu stößt man gelegentlich auf die Aussage, daß man systemtheoretisch zu nicht-reduktionistischen Erklärungen komme: ein Phänomen werde im Zusammenhang eines einbettenden Systems erklärt, dieses System sei aber hinwiederum durch das Phänomen mitbestimmt; d.h. man versteht das Ganze und das Einzelne in wechselseitiger Bedingtheit. Diese Perspektive ist nicht neu; sie findet sich schon in der philosophischen Hermeneutik und der Gestaltphilosophie/Gestaltpsychologie.

Meiner Meinung nach steckt hinter solchen Aussagen ein technisches Mißverständnis. Bei näherem Hinsehen findet man, daß systemtheoretische Modelle auf zwei Arten zur Erklärung von Phänomenen verwendet werden können. Die eine Art ist schlicht und einfach die gewohnte reduktionistische. Die andere Art ist zwar in gewissem Sinne nicht-reduktionistisch, sie beruht aber auf mathematisch klaren Universalitätseigenschaften dynamischer Systeme, denen nichts Geheimnisvolles anhaftet. Außerdem sind nicht-reduktionistische Erklärungen dieser Art nur "oberflächlich" und einer reduktionistischen Fundierung fähig.

Zur Erklärung betrachten wir die vier Felder in Abb. 6.

Im Feld 4 ist das zu erklärende, physikalische Systeme dargestellt. Feld 3 zeigt ein quantitatives systemtheoretisches Modell, das z.B. aus Differentialgleichungen bestehen wird. Zu ihrer Aufstellung ist eine genaue Kenntnis der systembestimmenden Variablen und deren dynamischer Zusammenhänge notwendig. In Feld 2 ist schematisch das Ergebnis einer qualitativen Analyse der Systemgleichungen dargestellt. Eine solche qualitative Analyse beschreibt den Zustandsraum durch Attraktoren, Bifurkationen und dergleichen. Im Feld 1 schließlich findet man die emergenten Phänomene, die im physikalischen System zu beobachten sind und die typischerweise in einem qualitativen Vokabular beschrieben werden.

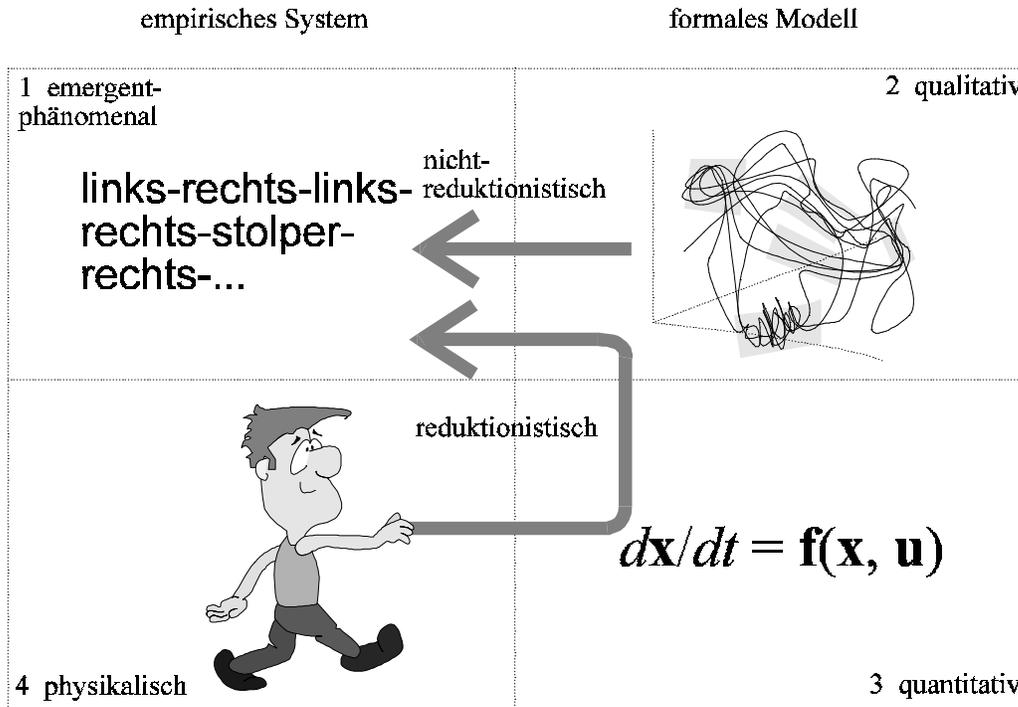


Abb. 6: Reduktionistische und nicht-reduktionistische systemtheoretische Erklärungen

Dieser Weg durch die Felder 4-3-2-1 entspricht einer reduktionistischen Erklärung. Das auf der qualitativen, emergenten Ebene beobachtete Verhalten des physikalischen Systems wird auf dem Umweg über ein formales Modell auf die niedrigere Ebene des physikalischen Substrats reduziert. Die Übergänge zwischen den Feldern sind alle nichttrivial. Der Übergang 4-3 beinhaltet die Aufgabe, geeignete Systemvariable zu identifizieren und deren dynamisches Verhalten zu erforschen. Beides kann beträchtliche Mühe bereiten (instruktives Beispiel einer solchen Suche nach den richtigen Variablen bei Clark et al. 1993). Der Übergang 3-2 ist die mathematische Analyse der Systemgleichungen. Wie in Kapitel 4 noch hervorgehoben wird, kann diese Aufgabe beliebig schwierig sein, selbst wenn man sich auf numerische Simulationen zurückzieht. Der Übergang 2-1 schließlich bedeutet eine Abbildung von formalen systemtheoretischen Konstrukten auf reale, qualitativ kategorisierte Phänomene. Hier ist Intuition gefragt, und die Zuordnung wird nie vollständig "hart" zu rechtfertigen sein.

Musterbeispiele solcher reduktionistischen Erklärungen findet man bei Yao & Freeman (1990) und Ewert & v. Seelen (1974).

Bei nicht-reduktionistischen Erklärungen erklärt man emergente Phänomene im realen System durch qualitative Eigenschaften von formalen Systemen. Man beschränkt sich also auf den Übergang 2-1. Diese Art der systemtheoretischen Erklärung ist außerhalb der Naturwis-

senschaften sehr verbreitet. Sie kann mehr oder weniger präzise durchgeführt werden. Ein Beispiel für eine recht präzise Durchführung ist die Deutung der spontanen Bewegungen eines Neugeborenen als chaotischer Attraktor bei Robertson et al. (1993). Hier wird die fraktale Dimension des Attraktors aus empirischen Daten berechnet. Das ist möglich, ohne ein formales Modell im Sinne des Feldes 3 zu besitzen, d.h. insbesondere, ohne sich darum zu kümmern, welches eigentlich die systembestimmenden Variablen sind. Aus der bloßen Tatsache, daß ein chaotischer Attraktor mit gewissen mathematischen Eigenschaften vorliegt, lassen sich nichttriviale Schlußfolgerungen (z.B. Stabilität, Dimensionalität eines Einbettungsraumes) ziehen.

Der Erklärungsgehalt solcher nicht-reduktionistischer Erklärungen beruht in der *Universalität* gewisser Eigenschaften dynamischer Systeme. Wenn *alle* Mitglieder in einer bestimmten Klasse von dynamischen Systemen eine bestimmte qualitative Eigenschaft haben, dann braucht man zur Erklärung eines emergenten Phänomens (in Feld 1) die genaue Form des *speziellen* dynamischen Modells (Feld 3) nicht zu kennen. Man muß nur belegen, daß das fragliche empirische System der betreffenden Klasse von dynamischen Systemen angehört, in denen bestimmte qualitative Phänomene (Feld 2) universal sind. Diese Art der Erklärung funktioniert deswegen so gut, weil die mathematische Forschung sehr interessante qualitative Eigenschaften in sehr großen Klassen dynamischer Systeme gefunden hat (z.B. Attraktortypen, Bifurkationstypen, Systemverhalten in der Nähe von Bifurkationen, Versklavungsprinzip, gekoppelte Oszillationen, Chaos).

Nicht-reduktionistische Erklärungen dieser Art sind nicht-tautologisch, da man aus der Erkenntnis, daß eine bestimmte qualitative Eigenschaft aus einer bestimmten Klasse dynamischer Systeme vorliegt, Voraussagen ableiten kann, die man dann wieder empirisch prüfen kann, was eine Falsifikation der ursprünglich angenommenen Erklärung möglich macht. Beispiele hierfür finden sich bei Wolff (1993).

Auch wenn solche nicht-reduktionistischen Erklärungen mithin echte Erklärungen sind, darf man nicht vergessen, daß sie im Prinzip um ein formales Modell (Feld 3) und eine Rückführung auf das physikalische Substrat des empirischen Systems (Feld 4) erweitert werden können, wodurch sie zu reduktionistischen Erklärungen werden und ihre Auflösung stark zunimmt. Nicht-reduktionistische Erklärungen sind Provisorien!

Ich vermute, daß in der populären bzw. außerwissenschaftlichen Diskussion um systemtheoretische Modelle häufig "nicht-reduktionistische Erklärung" mit "Rückkopplung" verwechselt wird. Die verschiedenen Teilsysteme und Systemvariable eines dynamischen Systems sind in der Regel auf zyklischen Interaktionspfaden miteinander rückgekoppelt. Daher kann man die Dynamik eines Teilsystems oder einer Variable nicht verstehen, ohne gleichzeitig die Dynamik der anderen mitzumodellieren. In diesem Sinne sind in rückgekoppelten Systemen die Teile nur mit dem Ganzen gemeinsam zu erklären. Diese Einsicht drückt jedoch nur eine charakteristische Randbedingung systemtheoretischer Erklärungen aus; sie stellt nicht schon ein Erklärungsprinzip selbst dar.

## 4 Diskussion

Im vorangegangenen Kapitel habe ich systemtheoretischer Methoden positiv gewürdigt. Die herausragenden Beiträge, die sie zum Verständnis kognitiver Prozesse liefern können, seien hier noch einmal vermerkt:

- Ein systemtheoretisches Modell eines Agenten kann gleichzeitig sowohl eine quantitative Beschreibungen des körperlich-mechanischen oder neuronalen Apparats beinhalten wie auch qualitative, emergente Phänomene erfassen, die zur Informationsverarbeitung des

Agenten gehören. In gewisser Weise wird so die scheinbare kategoriale Kluft zwischen Körper und Geist überbrückt. Diese Überbrückung schließt einen Übergang von einer kontinuierlichen, numerischen Beschreibungssprache zu einer diskreten, begrifflichen ein.

- Eine große Vielfalt zeitlicher Phänomene wird sichtbar und kann formal analysiert werden.
- Es wird erklärt, warum Entwicklungsprozesse sich in abwechselnde, relativ stabile Phasen und qualitative Sprünge gliedern. Der empirischen Entwicklungsforschung werden mathematische Techniken in die Hand gegeben, die eine viel feinere (insbesondere zeitliche) Modellierung und damit detailliertere Prädiktionen erlauben als die klassischen statistischen Beschreibungstechniken.
- Bei der Modellierung von Konzepten kann mit einigen klassischen Schwierigkeiten gut umgegangen werden. Sie betreffen z.B. Variabilität, Kontextsensibilität, Gestalteigenschaften, und terminologische Zyklen.

Es gibt jedoch beachtliche innere und äußere Hindernisse, die sich einer Verwendung systemtheoretischer Methoden in der Kognitionswissenschaft entgegenstellen.

Bei der **empirischen Analyse realer Systeme** geht es darum, aus verrauschtem Datenmaterial ein formales Modell zu konstruieren. Zwei vorbildliche Beispiele für solche Analysen sind die Modellierungen von periodischen Handbewegungen von Schönner et al. (1986) oder die Rekonstruktion eines chaotischen Attraktors als Modell der Spontanbewegungen von Neugeborenen (Robertson et al. 1993). So verschieden die beiden Arbeiten auch sind, so haben sie doch zweierlei gemeinsam: erstens ist das jeweils untersuchte System sehr einfach, und zweitens gelangen beide Forschungsteams mit den vorhandenen Techniken bereits an die Grenzen des derzeit Möglichen. Man kann vermuten, daß dieses an-Grenzen-Stoßen nicht nur auf einen ungenügenden Entwicklungsstand der verfügbaren Analysemethoden zu schieben ist, sondern auch auf eine prinzipielle Erkenntnisgrenze hinweist. Sie besteht darin, daß wirklich komplexe, kontinuierliche dynamische Systeme in ihrem qualitativen Verhalten typischerweise sensibel von einer Vielzahl von Parameterwerten abhängen. Solche Parameter in der erforderlichen Genauigkeit aus empirischen Daten zu rekonstruieren ist unmöglich, wenn die zur ausreichenden Parameterpräzisierung erforderliche Information größer ist als die Information, die praktisch aus empirischen, typischerweise sehr verrauschten Daten gewonnen werden kann. Einfacher gesagt: man kann nicht genau (oder lange) genug messen, um kleine, aber entscheidende Einflußgrößen in den Griff zu bekommen.

Diese Lage ist hinwiederum nicht hoffnungslos. Erstens kann man versuchen, informationsärmere, evtl. diskret approximierende Systemmodelle zu verwenden (eines meiner Motive für die Entwicklung der dynamischen Symbolsysteme, vgl. Jaeger 1994a,d). Zweitens kann man versuchen, das zu analysierende System in Teilsysteme zu zergliedern und diese einzeln zu rekonstruieren. Diese Taktik entspricht dem bewährten Standard der naturwissenschaftlichen Praxis. Allerdings muß man sehen, daß Teilsysteme in informationsverarbeitenden Systemen durch die Kopplungsinteraktion oft in ihrem qualitativen Verhalten verändert werden. Hier macht eine isolierte Untersuchung der Teilsysteme dann wenig Sinn (weitere Diskussion in Jaeger 1995b).

Tiefliedende scaling-up-Probleme finden sich auch bei der **mathematischen Analyse formaler Systeme**. Hier liegt ein mathematisches Modell des Systems bereits vor, und es geht um dessen qualitative Beschreibung (Attraktoren, Bifurkationen, Stabilitätsbereiche u.a.). Bekanntermassen sind die Lösungen von Systemen nichtlinearer Differentialgleichungen so gut wie nie analytisch angebar. Deswegen wird man rasch zu numerischen Simulationen genötigt. Aber nur in einfachen Fällen erhält man durch sie ein erschöpfendes Bild des Systemverhaltens. Bei komplexeren Systemen kann das Systemverhalten auch durch numerische Simulationen nurmehr punktuell für ausgewählte Randbedingungen exploriert werden. Beispiele sind die Untersu-

chungen des durch Differentialgleichungen modellierten Riechhirns von Yao & Freeman (1990) oder des sensomotorischen Kontrollsystems zur Trajektoriensteuerung bei Gaudiano & Grossberg (1991). Solche Explorationen sind keine geringe Leistung, aber sie laufen auf inspiriertes "Probefahren" hinaus.

Man kann also festhalten, daß systemtheoretische Beschreibungen des qualitativen Verhaltens empirischer oder formaler Systeme an **Komplexitätsgrenzen** stoßen. Der faszinierende Anspruch systemtheoretischer Methoden, komplexe Systeme als Ganze verständlich zu machen, kann de facto für intelligente Systeme nicht eingelöst werden. Die in der Physik und der Chemie mit so großem Erfolg untersuchten Systeme sind eben um Größenordnungen weniger komplex als jedes zur Intelligenzerzeugung ansetzende informationsverarbeitende System. Letztere Systeme bestehen typischerweise aus heterogenen Modulen, überspannen mehrere Granularitätsebenen, sind schnell und hochdimensional variierenden Randbedingungen ausgesetzt, wohingegen die "naturwissenschaftlichen" Systeme relativ homogen sind, höchstens zwei Granularitätsebenen umspannen (eine Mikroebene und eine emergente Makroebene), und typischerweise unter niedrigdimensionalen, langsam wechselnden Randbedingungen untersucht werden. Da schon in der Physik systemtheoretische Methoden rasch ausgereizt werden, kann es nicht verwundern, daß sie bei der Analyse intelligenter Systeme überfordert sind.

Es gibt schließlich **fundamentale Bedenken** gegenüber systemtheoretischen Methoden, die damit zu tun haben, daß mentale Prozesse möglicherweise kategorial verschieden sind von "bloß" physikalischen. Systemtheoretische Techniken sind für erstere entwickelt worden. Aslin (1993) diskutiert einige sich hier stellende Fragen.

Die heute von der Physik und Mathematik gelieferten mathematischen Methoden haben **keinen stabilen Stand** erreicht. Die mathematische und konzeptuelle Entwicklung in der naturwissenschaftlichen und mathematischen Theorie dynamischer Systeme ist in vollem Schwung. Wichtige Grundsatzfragen sind noch nicht richtig verstanden, und die Verwendung systemtheoretischer Methoden geschieht manchmal auf schwankem Grund. Hier sei beispielhaft auf das ungenügende Verständnis der Beziehungen zwischen Chaos, Noise und der Güte von Beobachtungsdaten hingewiesen (vgl. Preißl et al. 1990, Vassilicos et al. 1993, Millonas 1994). Zentrale Begriffe, darunter "Bifurkationen", "Chaos" und "Selbstorganisation" werden auch von Mathematikern ohne präzise Definition gebraucht. Offenbar gibt es sehr viele verschiedene Mechanismen der Selbstorganisation auf verschiedenen Zeitskalen, aber eine Klassifizierung ist nicht in Sicht. Diese Liste solcher Unvollkommenheiten ließe sich verlängern. Für die Kognitionswissenschaft bedeutet das, daß sie kein fertiges, in jeder Hinsicht verstandenes Instrumentarium geliefert bekommt. Es bedeutet aber auch die Freiheit -- eine wunderbare Herausforderung -- auf die eigenen Bedürfnisse zugeschnittene Methoden innerhalb der Systemtheorie zu entwickeln.

Erhebliche Hürden für die Verwendung systemtheoretischer Methoden in der Kognitionswissenschaft resultieren schließlich aus äußerlichen, **soziokulturellen Bedingungen**:

- Formale systemtheoretische Kenntnisse sind bei Kognitionswissenschaftlern noch kaum vorhanden. Da systemtheoretische Lehrbücher typischerweise einen mathematischen Hintergrund voraussetzen, wird die Assimilation systemtheoretischer Methoden nicht leicht fallen.
- Die Verwendung von logik-orientierten Techniken, die Computermetapher für kognitive Systeme, und die Hypothese der *physical symbol systems* macht einen guten Teil des Selbstbildes der Kognitionswissenschaft aus. Hieraus mag bei Vielen eine gewisse Abwehr

gegenüber systemtheoretischer Methoden resultieren, da sie mit jenen klassischen Grundeinstellungen nicht ohne weiteres vereinbar sind.

- Es besteht außerdem die Gefahr, daß systemtheoretische Ansätze den Mechanismen einer Modeerscheinung unterworfen werden. Einer kurzen Euphorie könnte aufgrund der vorher angedeuteten intrinsischen Schwierigkeiten eine so starke Ernüchterung folgen, daß auch langfristig produktive Ansätze vorzeitig wieder aufgegeben werden.

Damit möchte ich diesen Aufsatz beenden. Ich habe darin drei Ziele verfolgt. Erstens wollte ich dem Leser eine erste Orientierung im unübersichtlichen Gelände der Systemtheorie verschaffen. Zweitens wollte ich aufzeigen, welche enormen Erklärungspotentiale systemtheoretische Modellbildungen für die Kognitionswissenschaft haben. Und drittens wollte ich auf die nicht zu unterschätzenden Schwierigkeiten hinweisen, die einer raschen Einführung systemtheoretischer Methoden entgegenstehen.

Hoffentlich ist beim Leser am Ende die Antizipation der Potentiale stärker als der Respekt vor den Schwierigkeiten!

## Danksagung

Die Motivation zu dieser Arbeit erwuchs aus langen Gesprächen über dynamische Systeme und autonome Agenten, die ich mit Simone Strippgen geführt habe. Thomas Christaller und Luc Steels gewährten großzügige und entscheidende Unterstützung. Mannigfache Hinweise, Zuspruch und konstruktive Kritik kamen von Manfred Bierwisch, Reinhard Blutner, Bart de Boer, Tim van Gelder, Donal McKernan, Gregoire Nicolis, Robert Port, Emmet Spier, und Peter Stuer. Ihnen allen möchte ich herzlich danken. Die Arbeit wurde in einer ersten Version durch einen Werkvertrag von der GMD in St. Augustin ermöglicht; wesentliche Ergänzungen konnte ich dank eines Forschungsstipendiums der VUB Brüssel vornehmen.

## Literatur

- Abraham, F.D. (1990): A Visual Introduction to Dynamical Systems Theory for Psychology. Aerial Press, Santa Cruz
- Abraham, R. H., Shaw, Ch. D. (1992<sup>2</sup>): Dynamics: The Geometry of Behavior. Addison-Wesley, Redwood City 1992
- Arrowsmith, D.K., Place, C.M. (1992): Dynamic Systems - Differential Equations, Maps and Chaotic Behavior. Chapman and Hall, London 1992
- Aslin, Richard N. (1993): The Strange Attractiveness of Dynamic Systems to Development. In Smith & Thelen 1993a, 358-399
- Baayse, K., Dean, Th., Kaelbling, L.P. (1995): Learning Dynamics: System Identification for Perceptually Challenged Agents. Artificial Intelligence 72, 139-171
- Babloyantz, A., Lourenço, C. (1994): Computation with Chaos: A Paradigm for Cortical Activity. Proceedings of the National Academy of Sciences of the USA 91, 9027-9031
- Balkenius, Ch., Gärdenfors, P. (1991): Nonmonotonic Inferences in Neural Networks. In: Allen, J.A., Sandewall, E. (Hrsg.) (1991): Principles of Knowledge Representation and Reasoning. Proceedings of the 2nd International Workshop. Morgan Kaufman, San Mateo
- Barsalou, L.W. (1987): The Instability of Graded Structure: Implications for the Nature of Concepts. In: Neisser, U. (Hrsg.): Concepts and Conceptual Development: Ecological and Intellectual Factors in Categorization. Cambridge University Press, Cambridge 1987

- Barsalou, L.W. (1989): Intraconcept Similarity and its Implications for Interconcept Similarity. In: Vosniadou, S., Ortony, A. (Hrsg.): *Similarity and Analogical Reasoning*. Cambridge University Press, 76-121
- Bartlett, F.C. (1932): *Remembering: A Study in Experimental and Social Psychology*. Cambridge University Press, Cambridge 1932 (letzte Neuauflage 1977)
- Beer, R. (1995): A Dynamical Systems Perspective on Agent-Environment Interaction. *Artificial Intelligence* 72 (1/2), 173-216
- von Bertalanffy, Ludwig (1968): *General System Theory: Foundations, Development, Applications*. Brazillier, New York
- Bertenthal, B.I., Pinto, J. (1993): Complementary Processes in the Perception and Production of Human Movements. In: Smith & Thelen 1993a, 209-239
- Bickhard, M. (1993): Representational Content in Humans and Machines. *Journal of Experimental and Theoretical AI* 5 (1993), 285-333
- Braitenberg, V. (1978): Cell Assemblies in the Cerebral Cortex. In: Heim, R., Palm, G. (Hrsg.) (1978): *Theoretical Approaches to Complex Systems*. Springer Verlag, Berlin/Heidelberg/New York (Lecture Notes in Biomathematics 21), 171-188
- Brézillon, P. (1993) (Hrsg.): *Proceedings des IJCAI-93 Workshop "Using Knowledge in its Context"*. Technical Report LAFORIA 93/13, Institut Blaise Pascal, Université Paris VI et VII
- Briggs, J., Peat, D. (1989): *Turbulent Mirror*. Harper & Row, New York
- Brooks, R.A. (1991): New Approaches to Robotics. *Science* 253, 1227-1232
- Buhmann, J., Schulten, K. (1987): Storing Sequences of Biased Patterns in Neural Networks with Stochastic Dynamics. In: Eckmiller, R., von der Malsburg, C. (Hrsg.) (1987): *Neural Computers*. Proceedings des NATO Advanced Research Workshop on Neural Computers 1987. Springer Verlag, Berlin/Heidelberg/New York 1987, 231-242
- Campbell, G.J., Taylor, J.G. (1993): The Temporal Kohonen Map. *Neural Networks* 6, 441-445
- Carpenter, G.A., Grossberg, S. (1990): ART 3: Hierarchical Search Using Chemical Transmitters in Self-Organizing Pattern Recognition Architectures. *Neural Networks* 3 (2), 129-152
- Chalmers, D.J. (1990): Subsymbolic Computation and the Chinese Room. In: Dinsmore, J. (Hrsg.): *The Symbolic and Connectionist Paradigms: Closing the Gap*. Lawrence Erlbaum, Hillsdale N.J., 1992, 25-48
- Chalmers, D.J. (1992): Connectionism and Compositionality: Why Fodor and Pylyshyn Were Wrong. *Philosophical Psychology* 6, 1993, 305-319
- Clancey, W.J. (1993): Situated Action: A Neuropsychological Interpretation. *Cognitive Science* 17, No 1, 1993, 87-116
- Clark, J.E., Trully, T.L., Phillips, S.J. (1993): On the Development of Walking as a Limit-Cycle System. In: Smith & Thelen 1993a, 71-94
- Cruse, H., Bartling, Ch., Cymbalyuk, G., Dean, J., Dreifert, M. (1994): A Modular Artificial Neural Net for Controlling a Six-Legged Walking System. Proceedings der Konferenz "Prerational Intelligence in Robotics: from Sensorimotor Intelligence to Collective Behavior", Report Nr. 10 der Forschungsgruppe "Prerational Intelligence", Zentrum für interdisziplinäre Forschung (ZIF), Universität Bielefeld, 17-36
- Crutchfield, J.P. (1992): Semantics and Thermodynamics. In: Casdagli, M., Eubank, S. (eds.) (1992): *Nonlinear Modeling and Forecasting*. SFI Proceedings Vol. XII, Addison-Wesley, Redwood City etc., 317-359
- Crutchfield, J.P., Young, K. (1990): Computation at the Onset of Chaos. In: Zurek, W.H. (Hrsg.): *Complexity, Entropy, and the Physics of Information*. SFI Studies in the Sciences of Complexity, vol. VIII, Addison-Wesley, 1990, 223-269
- Dautenhahn, K., Christaller, Th. (1995): Remembering, Rehearsal and Empathy: Toward a Social and Embodied Cognitive Psychology for Artifacts. Proceedings des "Reaching for Mind" Workshop auf der AISB-95
- Davies, Paul (1987): *Cosmic Blueprint*. Deutsch: *Prinzip Chaos - die neue Ordnung des Kosmos*. Bertelsmann, München 1988
- Devlin, K. (1991): *Logic and Information*. Cambridge University Press, Cambridge
- Drescher, G.L. (1991): *Made-up Minds: A Constructivist Approach to Artificial Intelligence*. MIT Press, Cambridge, Mass.
- Dress, A., Hendrichs, H., Küppers, G. (Hrsg.) (1986): *Selbstorganisation - die Entstehung von Ordnung in Natur und Gesellschaft*. Piper, München/Zürich 1986
- Ebeling, W., Nicolis, G. (1992): Word Frequency and Entropy of Symbolic Sequences: a Dynamical Perspective. *Chaos, Solitons & Fractals* 2 (6), 635-650
- Eckhorn, R., Stoeker, M. (1994): Synthesizing Complex Perceptions I: Concepts of Visual Feature Associations Based on Neural Synchronization and Related Experimental Results in the Visual Cortex. Proceedings der Konferenz "Emergence of Prerational Intelligence in Biology: From Sensorimotor Intelligence

- to Collective Behavior", Part 1, Report Nr. 7 der Forschungsgruppe "Prerational Intelligence", Zentrum für interdisziplinäre Forschung (ZIF), Universität Bielefeld, 81-100
- Eigen, M., Schuster, P. (1977f): The Hypercycle: A Principle of Natural Self-Organization. *Naturwissenschaften* 64 (1977), 541-565 (Part A); 65 (1978), 7-41 (Part B); 65 (1978), 341-369 (Part C)
- Eigen, M., Winkler, R. (1975): *Das Spiel*. Piper, München
- Engel, A.K., König, P., Gray, C.M., Singer, W. (1990): Synchronization of Oscillatory Responses: A Mechanism for Stimulus-Dependent Assembly Formation in Cat Visual Cortex. In: Eckmiller, R., Hartmann, G., Hauske, G. (Hrsg.) (1990): *Parallel Processing in Neural Systems and Computers*. Elsevier/North Holland, 105-108
- Ewert, J.-P., von Seelen, W. (1974): Neurobiologie und System-Theorie eines visuellen Mustererkennungsmechanismus bei Kröten. *Kybernetik* 14 (1974), 167-183
- Flavell, J.H. (1968): *The Developmental Psychology of Jean Piaget*. Van Nostrand, Princeton 1968
- von Foerster, H. (1984): Erkenntnistheorien und Selbstorganisation. *DELFIN* 1984 (4), 6-19. Nachdruck in Schmidt, S.J. 1987a, 133-158
- Ford, Joseph (1989): What is Chaos, that We Should be Mindful of It? In: Davies, Paul (Hrsg.) (1989): *The New Physics*. Cambridge University Press, Cambridge, p. 348-371
- Gärdenfors, P. (1994): How Logic Emerges from the Flow of Information. In: van Eijck, J., Visser, A. (Hrsg.) (1994): *Foundations of Computation*. MIT Press, 49-77
- Gaudiano, P., Grossberg, S. (1991): Vector Associative Maps: Unsupervised Real-Time Error-Based Learning and Control of Movement Trajectories. *Neural Networks* 4 (2), 1991, 147-183
- van Geert, P. (1993): A Dynamic Systems Model of Cognitive Growth: Competition and Support Under Limited Resource Conditions. In: Smith & Thelen 1993a, 265-332
- van Gelder, T. (1995): Connectionism, Dynamics, and the Philosophy of Mind. Erscheint in den Proceedings der Konferenz "Philosophy and the Sciences of Mind. The Third Pittsburgh-Konstanz Colloquium in the Philosophy of Science", Konstanz, Mai 1995
- van Gelder, T., Port, R. (1994a): Beyond Symbolic: Toward a Kama-Sutra of Compositionality. In: Honavar, V., Uhr, L. (Hrsg.): *Artificial Intelligence and Neural Networks: Steps Toward Principled Integration*. Academic Press, 1994, 107-125
- van Gelder, T., Port, R. (1994b): It's About Time: An Overview of the Dynamical Approach to Cognition. Research Report 116, Cognitive Science Programme, Indiana University, Bloomington, Indiana. Erscheint auch als Vorwort zu van Gelder & Port (Hrsg.) (1995): *Mind as Motion: Explorations in the Dynamics of Cognition*. Bradford/MIT Press
- Giles, C.L., Omlin, C.W. (1995): Learning, Representation, and Synthesis of Discrete Dynamical Systems in Continuous Recurrent Neural Networks. Proc. of the IEEE Workshop on Architectures for Semiotic Modeling and Situation Analysis in Large Complex Systems. Erscheint in IEEE Press
- Goldberg, D.E. (1989): *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley, Reading, Mass. 1989
- Goschke, T., Koppelberg, D. (1990): Connectionist Representation, Semantic Compositionality, and the Instability of Concept Structure. *Psychological Research* 52, 253-270
- Gould, S.J. (1979): *Ever Since Darwin. Reflections in Natural History*. Norton, New York/London
- Greeno, J.G., Moore, J.L. (1993): Situativity and Symbols: Response to Vera and Simon. *Cognitive Science* 17, No 1, 1993, 49-60
- Grossberg, S., Somers, D. (1991): Synchronized Oscillations During Cooperative Feature Linking in a Cortical Model of Visual Perception. *Neural Networks* 4 (4), 1991, 453-466
- Haken, H. (1983): *Advanced Synergetics - Instability Hierarchies of Self-Organizing Systems and Devices*. Springer, Berlin/Heidelberg (Springer Series in Synergetics Vol. 20)
- Haken, H., Wunderlich, A. (1986): Synergetik: Prozesse der Selbstorganisation in der belebten und un belebten Natur. In Dress et al. 1986, 35-60
- Happel, B.L.M., Murre, J.M.J. (1995): Evolving Complex Dynamics in Modular Interactive Neural Networks. Preprint, Universität Leiden, eingereicht bei Neural Networks
- Harnad, S. (1990): The Symbol Grounding Problem. *Physica D* 42, 335-346
- Hart, J., Berndt, R.S., Caramazza, A. (1985): Category-specific Naming Deficit Following Cerebral Infarction. *Nature* 316, 439-440
- Harvey, I., Husbands, P., Cliff, D. (1994): Seeing the Light: Artificial Evolution, Real Vision. In: Cliff, D. et al. (Hrsg.) (1994): *From Animals to Animats III: Proc. der Third Int. Conf. on Simulation of Adaptive Behavior*. Bradford/MIT Press, 392-401
- Hasida, K. (1994): Dynamics of Symbol Systems. *New Generation Computing* 12 (1994), 285-310
- Hayashi, Y. (1994): Oscillatory Neural Networks and Learning of Continuously Transformed Patterns. *Neural Networks* 7 (2), 219-232

- Hayes, P.J., Ford, K.M., Agnew, N. (1994): On Babies and Bathwater. A Cautionary Tale. *AI Magazine*, Fall 1994, 15-26
- Hebb, D.O. (1949): *The Organization of Behavior*. Wiley, New York
- Hofstadter, D.R. (1979): *Gödel, Escher, Bach*. Basic Books, New York
- Hofstadter, D.R., Mitchell, M. (1993): The Copycat Project: A Model of Mental Fluidity and Analogy-Making. Erscheint in: Holyoak, K., Barnden, J. (Hrsg.): *Advances in Connectionist and Neural Computation Theory*, Vol. II: Analogical Connections. Ablex, Norwood, N.J. Leicht überarbeitet auch als Kap. 5 in Hofstadter, D. (1995): *Fluid Concepts and Creative Analogies*. Basic Books, New York
- Holland, J.H., Reitmann, J.S. (1978): *Cognitive Systems Based on Adaptive Algorithms*. In: Waterman, D.A., Hayes-Roth, F. (Hrsg.) (1978): *Pattern Directed Inference Systems*. Academic Press, New York, 313-329
- Jaeger, H. (1991): An Introduction to Dynamic Concept Systems. In: Boley, H., Richter, M.M. (Hrsg.), *Processing Declarative Knowledge. Proceedings of the PDK-91 at Kaiserslautern*, Springer Verlag, Berlin (Lecture Notes in Artificial Intelligence 567), 87-106
- Jaeger, H. (1992): A Type-free Semantics for Concepts in Contexts. In: Brézillon 1993, 51-61
- Jaeger, H. (1994a): *Dynamic Symbol Systems*. Dissertation, Technische Fakultät der Universität Bielefeld 1994
- Jaeger, H. (1994b): Gestaltbildung durch Informationsmaximierung: eine formale und algorithmische Rekonstruktion. In: Opwis, K. (Hrsg.) (1994): *Proceedings "Erste Fachtagung der deutschen kognitionswissenschaftlichen Gesellschaft"*, Psychologisches Institut, Universität Freiburg 1994, 32-34
- Jaeger, H. (1994c): On Modelling Behaviors and Concepts as Attractors. *Proceedings der Konferenz "On the Role of Dynamics and Representation in Adaptive Behaviour and Cognition (DRABC-94)"* in San Sebastian, Dezember 1994. Universidad del País Vasco, San Sebastian 1994, 171-173
- Jaeger, H. (1994d): An Introduction to Dynamic Symbol Systems. In: Hallam, J. (Hrsg.): *Hybrid Problems, Hybrid Solutions*. *Proceedings van de AISB-95*. IOS Press/Ohmsha, Amsterdam 1995, 109-120
- Jaeger, H. (1995a): *Dynamische Systeme in der KI und ihren Nachbarwissenschaften*. Arbeitspapiere der GMD 925, GMD, St. Augustin
- Jaeger, H. (1995b): *Modulated Modules: Designing Behaviors as Dynamical Systems*. Arbeitspapiere der GMD 927, GMD, St. Augustin
- Jaeger, H. (1995c): *Identification of Behaviors in an Agent's Phase Space*. Arbeitspapiere der GMD 951, GMD, St. Augustin
- Kaelbling, L.P. (1991): *Foundations of Learning in Autonomous Agents*. *Robotics and Autonomous Systems* 8 (1-2), 1991. Nachdruck in: Van der Velde, W. (Hrsg.) (1993): *Toward Learning Robots*. Bradford/MIT Press
- Katok, K., Hasselblatt, B. (1995): *Introduction to the Modern Theory of Dynamical Systems*. Cambridge University Press, Cambridge (*Encyclopedia of Mathematics and its Applications* 54)
- Kelso, J.A. Scott, Ding, Mingzhou, Schöner, Gregor (1993): *Dynamic Pattern Formation: A Primer*. In: Smith, Linda B., Thelen, Esther (Hrsg.) (1993): *A Dynamic Systems Approach to Development*. Bradford/MIT Press, Cambridge, Mass. 1993, 13-50
- Köhler, R. (1987): *System Theoretic Linguistics*. *Theoretical Linguistics* 14 (1987), 241-257
- Košecká, J., Bajcsy, R. (1994): *Discrete Event Systems for Autonomous Mobile Agents*. *Robotics & Autonomous Systems* 12 (1994), 187-198
- Krause, W. (1989): *Über menschliches Denken - Denken als Ordnungsbildung*. *Zeitschrift für Psychologie* 197, 1989, 1-30
- Kriz, Jürgen (1992): *Chaos und Struktur: Systemtheorie Band I*. Quintessenz Verlag, München
- Krohn, W., Küppers, G., Paslack, R. (1987): *Selbstorganisation - Zur Genese und Entwicklung einer wissenschaftlichen Revolution*. In S.J. Schmidt 1987a, 441-465
- Lakoff, G. (1987): *Women, Fire, and Dangerous Things*. The University of Chicago Press, Chicago
- Large, E.W., Kolen, J.F. (1994): *Resonance and the Perception of Musical Meter*. *Connection Science* 6(2&3), 177-208
- Lashley, K.S. (1951): *The Problem of Serial Order in Behavior*. In: Jeffress, L.A. (Hrsg.) (1951): *Cerebral Mechanisms in Behavior*, Wiley, 1951, 112-136
- Legendre, G., Miyata, Y., Smolensky, P. (1990a): *Harmonic Grammar - A formal multi-level connectionist theory of linguistic well-formedness: An application*. ICS Technical Report #90-4, University of Colorado at Boulder
- Legendre, G., Miyata, Y., Smolensky, P. (1990b): *Harmonic Grammar - A formal multi-level connectionist theory of linguistic well-formedness: Theoretical foundations*. ICS Technical Report #90-5, University of Colorado at Boulder
- Ling, C.X., Buchal, R. (1993): *Learning to Control Dynamic Systems with Automated Quantization*. In: Brazdil, P.B. (Hrsg.): *Machine Learning. Proceedings der ECML-93*. Springer Verlag, Berlin etc. (Lecture Notes in Artificial Intelligence), 372-377

- Mandelbrot, B.B. (1977): *The Fractal Geometry of Nature*. Freeman, New York 1977. Deutsch: *Die fraktale Geometrie der Natur*. Birkhäuser, Basel/Boston 1987
- Mangold-Allwinn, R. (1991): *Flexible Konzepte: Modelle, Experimente, Simulationen*. Habilitationsschrift, University of Mannheim, 1991
- Mani, D.R., Shastri, L. (1993): Reflexive Reasoning with Multiple Instantiation in a Connectionist Reasoning System with a Type Hierarchy. *Connection Science* 5, No 3/4 (1993), 205-242
- Matsuga, M., Yuille, A.L. (1994): Spatiotemporal Information Storage in a Content Addressable Memory Using Realistic Neurons. *Neural Networks* 7 (3), 1994, 419-440
- Maturana, H.R., Varela, F.J. (1984): *El árbol del conocimiento*. Deutsch: *Der Baum der Erkenntnis: die biologischen Wurzeln des menschlichen Erkennens*. Scherz, Bern/München 1987; Goldmann Verlag, 1992<sup>4</sup>
- Medin, D.L. (1989): Concepts and Conceptual Structure. *American Psychologist* 44 (12), 1469-1481
- Millonas, M. (1994): The Importance of Being Noisy. *The Bulletin of the Santa Fe Institute*, Summer 1994, 22-23
- Munro, P.W., Anderson, J.A. (1988): Tools for Connectionist Modeling: The Dynamical Systems Methodology. *Behavior Research Methods, Instruments, & Computers* 20(2), 1988, 276-281
- Nebel, B. (1990): Reasoning and Revision in Hybrid Representation Systems. *Lecture Notes in Artificial Intelligence* 422, Springer Verlag, Berlin/Heidelberg/New York
- Nebel, B. (1991): Terminological Cycles: Semantics and Computational Properties. In: Sowa, J. (Hrsg.) (1991). *Principles of Semantic Networks*. Morgan Kaufman, San Mateo, 331-361
- Nicolis, J.S., Tsuda, I. (1989): On the Parallel Between Zipf's Law and 1/f Processes in Chaotic Systems Possessing Coexisting Attractors. *Progress of Theoretical Physics* 82 (2), 254-274
- Nicolis, G. (1989): Physics of Far-from-equilibrium Systems and Self-Organization. In: Davies, Paul (Hrsg.) (1989): *The New Physics*. Cambridge University Press, Cambridge, p. 316-347
- Nyman, A. (1966): *Die Schulen der neueren Psychologie*. Verlag Hans Huber, Bern/Stuttgart 1966
- Paslack, R. (1989): »...da stellt ein Wort zur rechten Zeit sich ein«. *Die Karriere des Chaos zum Schlüsselbegriff*. Kursbuch 98, 121-140
- Popper, K.R., Eccles, J.C. (1977): *The Self and Its Brain - An Argument for Interactionism*. Springer Verlag, Heidelberg etc., 1977. Deutsch: *Das Ich und sein Gehirn*. Piper, München/Zürich 1982<sup>2</sup>
- Port, R.F., Cummins, F., McAuley, J.D. (1994): Naive Time, Temporal Patterns and Human Audition. *Research Report 118*, Cognitive Science Program der Indiana University at Bloomington, Indiana, 1994. Erscheint auch in van Gelder & Port 1995
- Preißl, H., Aertsen, A., Palm, G. (1990): Are Fractal Dimensions a Good Measure for Neural Activity? In: Eckmiller, R., Hartmann, G., Hauske, G. (Hrsg.) (1990): *Parallel Processing in Neural Systems and Computers*. North-Holland, Amsterdam 1990, 83-86
- Prigogine, I. (1980): *From Being to Becoming*. Freeman, San Francisco. Deutsch: *Vom Sein zum Werden*. Piper, München 1979
- Quillian, M.R. (1968): Semantic Memory. In: Minsky, M. (Hrsg.): *Semantic Information Processing*. MIT Press, Cambridge, Mass., 227-270
- Rechenberg, I. (1973): *Evolutionsstrategie: Optimierung technischer Systeme nach Prinzipien der biologischen Evolution*. Friedrich Frommann Verlag, Stuttgart-Bad Cannstadt
- Reiss, M., Taylor, J.G. (1991): Storing Temporal Sequences. *Neural Networks* 4(6), 773-781
- Richards, D.D. (1988): Dynamic Concepts and Functionality: The Influence of Multiple Representations and Environmental Constraints on Categorization. *Human Development* 31, 11-19
- Rickheit, G., Strohner, H. (1992): Towards a Cognitive Theory of Linguistic Coherence. *Theoretical Linguistics* 18 (2/3), 209-237
- Riedel, U., Kühn, R., van Hemmen, J.L. (1988): Temporal Sequences and Chaos in Neural Nets. *Physical Review A* 38(2), 1105-1108
- Rieger, B. (1985) (Hrsg.): *Dynamik in der Bedeutungskonstitution*. Buske, Hamburg
- Robertson, S.S., Cohen, A.H., Mayer-Kress, G. (1993): Behavioral Chaos: Behind the Metaphor. In: Smith & Thelen 1993a, 119-150
- Sacks, E. (1990): A Dynamics System Perspective on Qualitative Simulation. *Artificial Intelligence* 42 (1990), 349-362
- Sacks, E. (1991): Automatic Analysis of One-parameter Planar Ordinary Differential Equations by Intelligent Numerical Simulation. *Artificial Intelligence* 48 (1991), 27-56
- Scheier, Ch., Pfeifer, R. (1995): Classification as Sensory-Motor Coordination: A Case Study on Autonomous Agents. Erscheint in den Proceedings der ECAL-95
- Schmidt, S.J. (Hrsg.) (1987a): *Der Diskurs des Radikalen Konstruktivismus*. Suhrkamp, Frankfurt (stw 636)

- Schmidt, S.J. (Hrsg.) (1987b): Der Radikale Konstruktivismus: Ein neues Paradigma im interdisziplinären Diskurs. In Schmidt, S.J. (1987a), 11-88
- Schöner, G., Haken, H., Kelso, J.A.S. (1986): A Stochastic Theory of Phase Transitions in Human Hand Movement. *Biological Cybernetics* 53, 247-257
- Schweitzer, G., Wen, J. (1994): Where Neural Nets Make Sense in Robotics. In: Proceedings der Konferenz "Prerational Intelligence in Robotics: from Sensorimotor Intelligence to Collective Behavior", Report Nr. 10 der Forschungsgruppe "Prerational Intelligence", Zentrum für interdisziplinäre Forschung (ZIF), Universität Bielefeld, 109-135
- Seidenberg, M.S., Tanenhaus, M.K., Leiman, J.M., Bienkowski, M. (1982): Automatic Access of the Meaning of Ambiguous Words in Contexts: Some Limitations of Knowledge-Based Processing. *Cognitive Psychology* 14, 489-537
- Smith, E.E., Medin, D.L. (1981): *Categories and Concepts*. Harvard University Press, Cambridge, Mass.
- Smith, L.B., Thelen, E. (Hrsg.) (1993a): *A Dynamic Systems Approach to Development: Applications*. Bradford/MIT Press, Cambridge, Mass.
- Smith, L.B., Thelen, E. (1993b): From the Dynamics of Motor Skill to the Dynamics of Development. In: Smith & Thelen 1993a, 1-11
- Smith, L.B., Thelen, E. (1993c): Can Dynamic Systems Theory be Usefully Applied in Areas Other than Motor Development? In: Smith & Thelen 1993a, 151-170
- Smithers, T. (1994): What the Dynamics of Adaptive Behavior and Cognition Might Look Like in an Agent-Environment Interaction System. In: Proceedings des Workshops "On the Role of Dynamics and Representation in Adaptive Behavior and Cognition", Universidad del Pais Vasco, San Sebastian 1994, 135-153
- Smolensky, P. (1986): Information Processing in Dynamical Systems: Foundations of Harmony Theory. In: Rumelhart, D.E., McClelland, J.L. (Hrsg.): *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, Vol. 1, MIT Press, Cambridge, Mass. 1986, 194-281
- Stadler, M., Kruse, P. (1986): Gestalttheorie und Theorie der Selbstorganisation. *Gestalt Theory* 8, 75-98
- Stadler, M., Kruse, P. (1990): The Self-Organization Perspective in Cognition Research: Historical Remarks and New Experimental Approaches. In: Haken, H., Stadler, M. (Hrsg.) (1990): *Synergetics of Cognition*. Springer Series in Synergetics 45, Springer Verlag, 32-52
- Steels, L. (1993a): Building Agents out of Autonomous Behavior Systems. In: Steels, L., Brooks, R. (Hrsg.) (1993): *The "Artificial Life" Route to "Artificial Intelligence": Building Situated Embodied Agents*. Lawrence Erlbaum, New Haven 1993
- Steels, L. (1993b): A Mathematical Framework for Autonomous Robots. Proc. der EWAIC 93, Moskau 1993, 333-334
- Steels, L. (1994a): Mathematical Analysis of Behavior Systems. In: Gaussier, P., Nicoud, J.-D. (Hrsg.): *Proceedings of the "From Perception to Action Conference"*, Lausanne, Sept. 1994, IEEE Computer Society Press, Los Alamitos, CA, 88-95
- Steels, L. (1994b): Emergent Functionality in Robotic Agents through On-Line Evolution. In Brooks, R., Maes, P. (Hrsg.) (1994): *Artificial Life IV*. Proceedings des Fourth Int. WS on the Synthesis and Simulation of Living Systems. Bradford/MIT Press, 8-14
- Steinkühler, U., Cruse, H. (1994): A Holistic Model for the Control of an Arm with Redundant Degrees of Freedom. In: Proceedings der Konferenz "Prerational Intelligence in Robotics: from Sensorimotor Intelligence to Collective Behavior", Report Nr. 10 der Forschungsgruppe "Prerational Intelligence", Zentrum für interdisziplinäre Forschung (ZIF), Universität Bielefeld, 149-163
- Stengel, R.F. (1986): *Stochastic Optimal Control*. Wiley, New York
- Stoecker, M., Eckhorn, R. (1994): Synthesizing Complex Perceptions II: Models of Visual Feature Associations Based on Neural Synchronization. Proceedings der Konferenz "Emergence of Prerational Intelligence in Biology: From Sensorimotor Intelligence to Collective Behavior", Part 1, Report Nr. 7 der Forschungsgruppe "Prerational Intelligence", Zentrum für interdisziplinäre Forschung (ZIF), Universität Bielefeld, 101-114
- Suchman, L.A. (1987): *Plans and Situated Actions: The Problem of Human-Machine Communication*. Cambridge University Press, Cambridge, Mass., 1987
- Tani, J. (1995): Embedding a Grammatical Description in Deterministic Chaos: An Experiment in Recurrent Neural Learning. Erscheint in *Biological Cybernetics*
- Tino, P., Horne, B.G., Giles, C.L. (1995): Finite State Machines and Recurrent Neural Networks - Automata and Dynamical Systems Approach. Technical Report, UMIACS-TR-95-1 und CS-TR-3396, Institute for Advanced Computer Studies, University of Maryland
- Treisman, A. (1986): Properties, Parts, and Objects. In: Boff, K. R., Kaufman, L., und Thomas, J.P. (Hrsg.): *Handbook of Perception and Human Performance*, vol. 2. John Wiley and Sons, New York 1986, Kapitel 35

- Triggs, B. (1994): Model-based Sonar Localisation for Mobile Robots. *Robotics & Autonomous Systems* 12, 173-186
- Tucker, M., Hirsh-Pasek, K. (1993): Systems and Language: Implications for Acquisition. In Smith & Thelen 1993a, 359-384
- Vassilicos, J.C., Demos, A., Tata, F. (1993): No Evidence of Chaos But Some Evidence of Multifractals in the Foreign Exchange and the Stock Markets. In: Crilly, A.J., Earnshaw, R.A., Jones, H. (Hrsg.)(1993): *Applications of Fractals and Chaos*. Springer Verlag, New York/Berlin/Heidelberg, 249-265
- Vera A.H., Simon, H.A. (1993a): Situated Action: A Symbolic Interpretation. *Cognitive Science* 17 No. 1, 1993, 7-48
- Vera A.H., Simon, H.A. (1993b): Situated Action: Reply to William Clancey. *Cognitive Science* 17 No. 1, 1993, 117-135
- Verschure, P.F.M.J. (1993): The Cognitive Development of an Autonomous Behaving Artifact: The Self-Organization of Categorization, Sequencing, and Chunking. *Proceedings der Konferenz "Prerational Intelligence - Phenomenology of Complexity Emerging in Systems of Agents Using Simple Rules"*, Report Nr. 2 der Forschungsgruppe "Prerational Intelligence", Zentrum für interdisziplinäre Forschung (ZIF), Universität Bielefeld, 95-120
- Wagner, G. (1986): Evolution der Evolutionsfähigkeit. In Dress et al. 1986, 121-147
- Waltz, D. L., Pollack, J.B. (1985): Massively Parallel Parsing: A Strongly Interactive Model of Natural Language Interpretation. *Cognitive Science* 9, 51-74.
- Weidemann, H.-J., Pfeiffer, F. (1994): The Control System of the Six-Legged TUM Walking Robot. In: *Proceedings der Konferenz "Prerational Intelligence in Robotics: from Sensorimotor Intelligence to Collective Behavior"*, Report Nr. 10 der Forschungsgruppe "Prerational Intelligence", Zentrum für interdisziplinäre Forschung (ZIF), Universität Bielefeld, 165-168
- von Weizsäcker, E.U. (1987): Brückenkonzepte zwischen Natur- und Geisteswissenschaften: Selbstorganisation, Offene Systeme und Evolution. *Sozial-ökologische Arbeitspapiere* 17, Forschungsgruppe Soziale Ökologie, Frankfurt/Main 1987, 1-27
- Wiener, N. (1961<sup>2</sup>): *Cybernetics*. MIT Press, Cambridge, Mass.
- Wildgen, W. (1985): *Archetypensemantik. Grundlagen für eine dynamische Semantik auf der Basis der Katastrophentheorie*. Gunter Narr Verlag, Tübingen
- Wilks, Y. (1975): A Preferential, Pattern-Seeking Semantics for Natural Language Inference. *Artificial Intelligence* 6, 1975, 53-74
- Willke, Helmut (1987): *Systemtheorie*, 2. Auflage. Gustav Fischer Verlag, Stuttgart 1987 (UTB 1161)
- Wolff, P.H. (1993): Behavioral and Emotional States in Infancy: A Dynamic Perspective. In Smith & Thelen 1993a, 189-208
- Woolridge, M., Jennings, N.R. (1995): Intelligent Agents: Theory and Practice. *The Knowledge Engineering Review* 10 (2), 115-152
- Wunsch, Gerhard (1985): *Geschichte der Systemtheorie*. Oldenbourg, München 1985
- Yamauchi, B., Beer, R. (1994): Integrating Reactive, Sequential, and Learning Behavior Using Dynamical Neural Networks. In: Cliff, D. et al. (Hrsg.) (1994): *From Animals to Animats III. Proc. der dritten internationalen Konferenz "On the Simulation of Adaptive Behavior"*, Bradford/MIT Press, 382-391
- Yao, Y., Freeman, W.J. (1990): A Model of Biological Pattern Recognition with Spatially Chaotic Dynamics. *Neural Networks* 3, No. 2, 153-170
- Zadeh, L.A. (1969): The Concept of System, Aggregate, and State in System Theory. In Zadeh & Polak 1969, 3-42
- Zadeh, L.A., Polak, E. (Hrsg.) (1969): *System Theory*. McGraw-Hill, New York 1969 (Inter-University Electronics Series Vol. 8)
- Zalama, E., Gaudiano, P., Coronado, J.L. (1995): A Real-Time, Unsupervised Neural Network for the Low-Level Control of a Mobile Robot in a Nonstationary Environment. *Neural Networks* 8 (1), 103-123
- Zwölfer, H. (1986): Insektenkomplexe an Disteln - ein Modell für die Selbstorganisation ökologischer Kleinsysteme. In: Dress et al. 1986, 181-218