

Können wir das Problem der Echtzeitkognition lösen?

Alexander Riegler

CLEA, Freie Universität Brüssel
Rue de la Strategie 33, B-1160 Brüssel, Belgien
Email: ariegler@vub.ac.be

Abstract. Eines der größten Probleme in der Artificial Life ist die Frage, wie Lebewesen mit dem "Information Overload" zurecht kommen, d.h. wie sie aus der Flut der Umweltreize die für sie relevante Information in Echtzeit herausfiltern. Ausgeklügelte informationsverarbeitende Mechanismen wurden konstruiert, die das Problem aber nur noch verschlimmern: Wie können Lebewesen mit ihrem in der Regel begrenzten kognitiven Fähigkeiten jemals eine derartige computationale Maschinerie beherbergen? Einem Vorschlag Karl Poppers folgend möchte ich für eine alternative Perspektive argumentieren und daraus generelle Richtlinien für künstliche Systeme darlegen.

Einleitung

Es war einmal ein kleiner Roboter, so beginnt eine bekannte Analogie von Daniel Dennett (1984), dessen Ersatzbatterie zusammen mit einer Bombe in einem Raum untergebracht war. Der Roboter sah, daß sich die Batterie auf einem Wagen befand, den er bloß aus dem Zimmer zu schieben hatte, um die Batterie zu retten. Da sich die Bombe ebenfalls auf dem Wagen befand, bewahrte die Handlung weder Batterie noch Roboter vor dem Untergang. Zwar hatte der Roboter erkannt, daß auch die Bombe auf dem Wagen lag, konnte aber nicht schlußfolgern, daß das Schieben des Wagens auch gleichzeitig die Bombe mitbeförderte. Ein des Schlußfolgerns mächtiger Nachfolgeroboter wurde konstruiert und mit demselben Problem konfrontiert. Nachdem er Batterie und Bombe auf dem Wagen sah, verharrte er in Bewegungslosigkeit, bis die Bombe explodierte: zu sehr war er mit allen möglichen Implikationen des Schiebens beschäftigt, so daß er nicht rasch genug handeln konnte. Er wußte einfach nicht, daß es müßig ist, die Auswirkungen des Schiebens auf die Farbe der Wände in dem Raum zu untersuchen. Ein weiteres Nachfolgemodell sollte also diese unsinnigen von den sinnvollen Auswirkungen unterscheiden können. Aber auch dieser Roboter verweilte regungslos bis zur Explosion, da er mit dem Abwägen "sinnvoll oder nicht?" beschäftigt war.

Keine dieser künstlichen Wesen war fähig, die brenzlige Situation in Echtzeit zu lösen. Zuviele Dinge sind zu bedenken, zuviel Information abzuwägen, um zur Handlung schreiten zu können. Wenn aber Artefakte jemals etwas "Lebendiges" sein wollen, dann müssen sie diese Schwierigkeiten meistern können. Das betrifft Roboter im gleichen Ausmaß wie "Software-Wesen".

Ich möchte daher in dieser Arbeit auf das Problem der, wie ich es nenne, "Echtzeitkognition" genauer eingehen und untersuchen, weshalb natürliche Geschöpfe gegenüber den Dennettschen Robotern enorme Geschwindigkeitsvorteile genießen. Bei diesem Vorhaben ist die Integration von Konzepten aus vielen Bereichen recht hilfreich: Ethologie, Konstruktivismus, Wissenschaftstheorie und Kognitionspsychologie.

Darauf aufbauend stelle ich einige grundlegende Richtlinien für künstliche kognitive Systeme auf, um diesem, wie es auch oft genannt wird, "Information Overload" zu begegnen.

Das Problem der Echtzeitkognition

Der erste Roboter scheitert, weil er die Implikationen seines Handelns nicht vorhersehen kann und dadurch eine "naive" Reaktion an den Tag legt. Will ein Roboter aber alle möglichen Implikationen seines Handelns berücksichtigen, so verharrt er in endlosem Berechnen, das keine Zeit für Handlungen gewährt. Das liegt daran, daß es eine beliebige große Anzahl möglicher Auswirkungen einer einzelnen Aktion gibt, wovon die meisten allerdings letztendlich irrelevant sind. Die Hoffnung besteht, daß die verbleibenden relevanten Implikationen nur von geringer Anzahl sind und deshalb rasches Handeln erlauben. Das Problem des Auseinanderhaltens von Sinnvollem und Nicht-Sinnvollem erfordert aber ebenso einen schier endlosen Berechnungsaufwand, der wiederum keine Zeit zum Reagieren zuläßt.

Viele Autoren, insbesondere Kritiker der Artificial Intelligence, argumentieren hier zu recht, daß das Frame-Problem eine Konsequenz des Logik-orientierten Zugangs zu Artificial Intelligence und Artificial Life ist. Das bedeutet, daß Sachverhalte in Einzelaussagen (Propositionen) zerlegt werden, um danach – ähnlich dem deduktiven Verfahren in der Mathematik – Schlüsse ziehen zu können. Wie das Beispiel zeigt, kann es so in tierischen Lebewesen, und, wie ich später argumentieren werde, im menschlichen Denken, nicht funktionieren. Tiere (wie auch Menschen) haben weitaus begrenzte Berechnungskapazitäten was das Ziehen von Schlüssen betrifft, dennoch sind sie in der Lage, mit komplizierten Situationen fertig zu werden. Ihr Überleben im Verlauf der evolutionären Selektion ist ein Beweis dafür.

Wie sieht nun eine Lösung für das Problem der Echtzeitkognition aus? Wie machen es natürliche Lebewesen, und können wir dieses Prinzip für künstliche Wesen, eben Roboter, abschauen? Auf den ersten Blick scheint der Implikationen-berechnende Roboter aus der Geschichte einem "Information Overload" gegenüber zu stehen. Dieser Begriff sagt, daß das betreffende Wesen mit einem Ausmaß an Information konfrontiert ist, das seine kognitiven Fähigkeiten überschreitet. Das obige Beispiel ist zwar (für philosophische Zwecke) konstruiert, dennoch läßt es erahnen, daß es in der "realen Welt" noch weitaus komplizierter zugeht, daß die Informationsfülle, die uns täglich begegnet, enorm sein muß. Und Lebewesen sind angesichts der unendlichen Informationsfülle ihrer Umgebung *stets* einem Information Overload ausgesetzt. Ein kleines mathematisches Beispiel enthüllt rasch die Komplexität der Entscheidungsfindung: Nehmen wir an, daß ein Wesen für jeden Entscheidungspunkt in einer Sequenz von $n = 10$ Handlungen aus einem Repertoire von $k = 10$ Handlungen wählen kann, um eine Lösung für ein Problem zu erreichen. Dies führt zu einer kombinatorischen Explosion von $k^n = 10^{10}$ Möglichkeiten!¹ Wie kann man unter solchen Bedingungen *jemals* zu einer Lösung gelangen?

Die Antwort, die ich hier vorstelle, muß daher lauten: Kognition besteht nicht darin, Information oder Fakten abzuarbeiten, sondern darin, diese (in geeigneter Weise) zu ignorieren.

Sehen wir uns zunächst genauer an, was ich hier mit Komplexität meine.

1. Die kombinatorische Explosion von Entscheidungen macht es auch sehr schwierig einen unschlagbaren Schachcomputer zu konstruieren, da dieser, im Gegensatz zu einem menschlichen Meisterspieler, alle Möglichkeiten durchprobieren muß, um den günstigsten Zug auswählen zu können.

Komplexität

Heinz von Foerster (1990) liefert eine nützliche Definition für potentielle Komplexität, indem er zwischen trivialen und nicht-trivialen Maschinen unterscheidet.

Eine *triviale Maschine* ist eine Maschine deren Arbeitsschritte nicht von vorangegangenen Schritten abhängig ist. Sie kann durch eine Funktion p beschrieben werden, die jede Eingangsvariable x auf eine Ausgangsvariable y abbildet: $p(x) \rightarrow y$. Wenn man das Verhalten einer derartigen Maschine lange genug beobachtet, so läßt sich p verhältnismäßig leicht daraus bestimmen. Das *Problem der Identifikation* ist damit lösbar: triviale Maschinen sind analytisch determinierbar und voraussagbar. Jeder Getränkeautomat ist ein typisches Exemplar einer trivialen Maschine: Der Druck auf eine bestimmte Auswahl Taste liefert stets ein Getränk derselben Marke unabhängig davon, was der vorherige Benutzer gewählt hat.

Im Gegensatz dazu weisen *nicht-triviale Maschinen*² einen internen Zustand z auf und umfassen zwei Funktionen: (1) Eine Effektfunktion p_z die ein zustandsabhängiges Abbilden von x nach y implementiert: $p_z(x) \rightarrow y$. Das bedeutet, daß sie die Eingangsvariable gemäß des eigenen internen Zustandes behandelt. (2) Eine Zustandsübergangsfunktion p_x , die den internen Zustand gemäß der Eingangsvariable ändert: $p_x(z) \rightarrow z'$. Mit anderen Worten: Gleicher Input wird unterschiedlich verarbeitet, abhängig vom internen Zustand, der sich wiederum aufgrund des Inputs ändert. Die Arbeitsweise einer nicht-trivialen Maschine hängt damit von ihrer eigenen Geschichte ab.

Für nicht-triviale Maschinen läßt sich das Problem der Identifikation im generellen Fall plötzlich nicht mehr lösen. Warum ist das so? Betrachten wir eine Maschine, die zwei verschiedene interne Zustände annehmen kann und jeweils vier Eingangs- und vier Ausgangsvariablen akzeptiert. Die Anzahl verschiedener Modelle, die eine derartig einfache Maschine beschreiben, ist $4^4 \cdot 4^4 = 2^{16} = 65536$. Erweitern wir die Anzahl der internen Zustände um eins, so existieren bereits $2^{24} > 16$ Millionen Modelle, die deren Funktionieren beschreiben. Lassen wir den Analytiker der Maschine im unklaren, wieviele Zustände, Ein- und Ausgangsgrößen die Maschine hat, so erhöht sich die Anzahl möglicher Modelle auf etwa 10^{155} : das System wird zur "Black Box". Eine derartige große Zahl von Modellen hätte selbst der denkbar leistungsfähigste Computer seit Bestehen des Planeten Erde nicht durchrechnen können, denn sie übersteigt das sogenannte *Bremermann-Limit* von 10^{93} bits.³

Biologisch betrachtet läßt sich die Unterscheidung zwischen der Effekt- und der Zustandsfunktion mit der Arbeit eines Ethologen bzw. mit der eines Physiologen in Verbindung bringen. Der Ethologe beobachtet, welche Reaktionen y ein Tier unter gewissen Umweltbedingungen x zeigt. Er protokolliert die Abbildungen von x nach y , also die Effektfunktion. Er kann aber nur Mutmaßungen über die Zustandsfunktion anstellen⁴, in welchem Umfang sich der interne Zustand des Tieres (etwa der Motivation) ändert. Die Physiologin hat gute Kenntnis über die Substanz der internen Zustände, weiß also, wie ein gegebener Input eine physiologische Größe ändert. Sie kann aber daraus nicht auf das Verhalten des Tieres schließen.

Dieser Sachverhalt der ungemein rasch wachsenden Komplexität muß auch in Verbindung mit dem "*evolutionären Schneeballeffekt*" (Clark 1997) gesehen werden. Er besagt, daß je komplexer ein Lebewesen (oder Entität generell) ist, desto abhängiger ist es von seiner Geschichte. Die Beobachtung, daß ein Einzeller stets dasselbe Verhalten beim Ein-

2. In der Informatik auch als Turing-Maschinen bezeichnet.

3. Hans Bremermann (1962) stellte fest daß "[n]o data processing system, whether artificial or living, can process more than $2 \cdot 10^{47}$ bits per second per gram of its mass".

4. Genauer gesagt: Der Ethologe kann lediglich von der eigenen anthropomorphen Perspektive auf das Lebewesen projizieren. Im alltäglichen Leben machen wir ähnliches, wenn wir beispielsweise einem schwanzwedelndem Hund die Empfindung der Freude unterstellen.

treten eines gewissen Umstandes zeigt, während unser eigenes menschliches Verhalten sehr mit unseren Erinnerungen (erfaßt in "internen Zuständen") zusammenhängt, bestätigt diese Einsicht. Wir haben es bei kognitiv entwickelten Wesen, und das könnten letztendlich auch Artefakte einmal werden, mit sehr geschichtsabhängigen und damit extrem komplexen Geschöpfen zu tun.

Jeder kognitive Akt scheint damit praktisch unmöglich zu sein. Einen möglichen Ausweg aus diese offenbar hoffnungslosen Situation der unbewältigbaren Komplexität der Realität wird durch die Wissenschaftstheorie Karl Poppers zum Thema "Denken" vorgezeichnet.

Denken nach Popper

Popper (1979) schlägt zwei konträre Analogien vor, die den Vorgang der Kognition kennzeichnen.

Seine *Eimertheorie des Denkens* schließt an die Ansicht "Nihil in intellectu quod non prius in sensu" an, also an die Meinung, daß wir nicht über etwas denken können, das uns nicht zuvor durch unsere Sinne erreicht hat. Diese Analogie vergleicht Kognition mit einem Eimer, der sukzessiv mit Wissen durch unsere Sinnesorgane gefüllt wird. Wie aber wird dieses Wissen repräsentiert (Peschl und Riegler 1999), in welcher "Sprache" wird es formuliert? Wie Forschung auf dem Gebiet der Artificial Intelligence und deren Kritik gezeigt haben (z.B. Dreyfus und Dreyfus 1990), enden wir mit einem solchen Ansatz in genau denselben Schwierigkeiten, die auch den Robotern in der Einführung zu schaffen machen. Es sind dies essentiell zwei Probleme: (1) Das *Frame-Problem*, das die Angabe eines wirkungsvollen Formalismus verlangt, der beschreibt, was sich in der Umwelt eines Wesens ändert und was konstant bleibt. Fehlt der Formalismus, so führt die Beschreibung von Handlungen und deren Konsequenzen zu der oben genannten kombinatorischen Explosion. (2) Das *Symbol Grounding-Problem* (Harnad 1990), das sich auf die Frage nach der Bedeutung einer "Wissenssprache" bezieht: Wie kann die Bedeutung von definitionsgemäß bedeutungsfreien Symbolen – die allein aufgrund ihren formalen Eigenschaften verarbeitet werden – in etwas anderem als in anderen bedeutungsfreien Symbolen verankert sein?

Die *Scheinwerfertheorie des Denkens* geht einen umgekehrten Weg. Wissen wird aktiv in Form von zunächst ungeprüften Vermutungen *konstruiert*. Erfahrung ist damit *nicht* die Quelle dieser Hypothesen. Erst Beobachtung wählt aus diesen (widersprüchlichen) Vermutungen die "brauchbarste" aus. Kognition geht also von einem zuvor gezimmerten "Weltbild" aus in Richtung Erwartungen, die einer Bestätigung bedürfen. Solange Erwartungen mit den sensorischen Bestätigungen übereinstimmen, kann das zur Zeit bestehende Hypothesengebäude aufrecht erhalten bleiben. Diese Vorstellung ist ähnlich der Situation, die man in einer tiefschwarzen Nacht bei Stromausfall in der eigenen Wohnung vorfindet. Im allgemeinen hat man trotz der fehlenden Beleuchtung keine Probleme, sich in den eigenen vier Wänden zurechtzufinden. Das Wissen über räumliche Zusammenhänge helfen bei der Orientierung, unterstützt durch gelegentliches Herumtasten, um die Richtigkeit der momentanen "Vermutung" zu überprüfen.

Kann die Scheinwerfertheorie gerechtfertigt werden? Arbeitet unser Denken nach der Scheinwerfermethode, die Erfahrung sekundär einstuft, so stellt sich die Frage, ob wir damit nicht von einer Tabula Rasa starten müssen, wir also anzunehmen haben, daß wir zu Beginn kein Wissen über die Umwelt haben. Tatsächlich ist dies der Fall, wenn wir uns vergegenwärtigen, daß unser Nervensystem keinerlei Information über die Bedeutung einer übertragenen Nachricht liefert. Dieses Prinzip der indifferenten Kodierung verlangt geradezu die Annahme, daß das Gehirn erst "Sinn" erzeugt. Damit ist unser Denken dadurch charakterisiert, daß es Gebäude von Vermutungen verwendet, die diesen anonymen Nervensignalen entspringen. In weiterer Konsequenz ist damit die Auffassung einer Kor-

responzenztheorie überflüssig, nämlich, daß Kognition Information von der Umgebung abbildet (Peschl und Riegler 1999). Vielmehr liegt eine (auf die Informationsverarbeitung bezogene) *operationale Geschlossenheit* (Maturana 1982) vor: Jede Zustandsänderung der relativen Aktivität einer Neuronengruppe führt zu einer Zustandsänderung der relativen Aktivität dieser oder einer anderen Neuronengruppe (Winograd und Flores 1986). Sinnesreize stellen dabei nur Perturbationen dar, welche den kognitiven Apparat in seinem Operieren zwar beeinflussen, aber nicht determinieren. Nicht "Information von Außen" charakterisiert die Funktionsweise des kognitiven Apparates, sondern die fortdauernde interne Konstruktion der Welt, die ankommende Perturbationen lediglich zu interpretieren sucht. Die Entwicklung vom Neugeborenen zum Erwachsenen im speziellen und die evolutionäre Entwicklung kognitiver Kompetenz im allgemeinen spiegelt diese fortdauernde Konstruktion wider.

Ein in diesem Zusammenhang vielfach zitiertes Beispiel ist das des Unterseebootnavigators, der sich (in Anbetracht der undurchdringlichen Finsternis in großen Tiefen) völlig auf Anzeiginstrumente (also indifferente Nervensignale) verläßt und in Abhängigkeit davon gegebenenfalls diverse Hebeln und Knöpfe bedient. Aus unserer Perspektive zeigen die Instrumente selbstverständlich "etwas draußen" an, aber das ist für das korrekte Navigieren irrelevant. Worauf es ankommt, ist das Aufrechterhalten von Relationen zwischen Instrumenten. Popper belegt diesen Umstand mit der Aussage: "Observations are secondary to hypotheses" (1979).

Da wir unser Problem der Echtzeitkognition auch für künstliche Geschöpfe lösen wollen, müssen wir uns fragen, ob nur das menschliche Gehirn in der Lage ist, aus den eintreffenden indifferenten Nervensignalen Sinn zu machen, oder ob das Prinzip der operativen Geschlossenheit und damit die Scheinwerfertheorie auch in einem weiteren Rahmen gültig ist.

Kanalisation von Verhalten

Die kognitiven Leistungen von Tieren sind nicht direkt erschließbar, sondern müssen über den Umweg des an den Tag gelegten Verhaltens erschlossen werden. Eine der ergiebigsten Quellen dafür sind die Arbeiten von Konrad Lorenz. Seine bekannten Beschreibungen vom Eirückholverhalten der Graugans (Lorenz und Tinbergen 1939) liefern ein paradigmatisches Bild, das zahlreiche Aspekte tierischer Kognition, wie etwa dessen Regelmäßigkeit, darlegt. Bemerkte eine brütende Graugans das Herausfallen eines Eies aus dem Nest, so versucht sie mit der Unterseite ihres Schnabels das Ei wieder ins Nest zurückzurollen. Interessanterweise bricht das Tier seine Bewegung nicht ab, wenn das Ei auf halber Strecke entfernt wird. Erst wenn der Nestrand erreicht ist, wird das Verhaltensmuster beendet. Damit verhält sich die Graugans nach einer allgemein formulierbaren Regel: Wenn ein Schlüsselreiz gegeben ist, dann starte eine bestimmte Verhaltenssequenz. Der Reiz kann dabei durchaus komplexer (zusammengesetzter) Natur sein. Analog kann die Verhaltenssequenz ebenfalls kompliziert sein: Während die Graugans das Ei zurückrollt, versucht sie Unebenheiten des Bodens mit übersteuernden Bewegungen des Schnabels in die entsprechende Richtung auszugleichen, um damit das rollende Ei auf einem geraden Kurs zu halten. In unserer Terminologie ausgedrückt, hat das Tier die Vermutung, daß das Ei gerade zurückrollt, holt aber laufend Bestätigungen dafür ein, um ggf. korrigieren zu können.

Mit dem regelhaften Aspekt des Verhaltens ist eine gewisse *Erwartungsgetriebenheit* verbunden: Im Falle der Graugans unterbricht das Entfernen des Eies das Verhalten nicht. Das deutet darauf hin, daß es während der Handlung zu einer Reduzierung der Verarbeitung sensorischer Reize kommt. Da normalerweise kein Ethologe das Ei entfernt, macht dieses Ignorieren in evolutionärer Perspektive Sinn. Weiters ist trotz dieser und anderer Einschränkungen das Verhalten keineswegs starr, sondern weist kognitive Flexibilität auf, die sich in der fortlaufenden Kompensation für Seitwärtsrollen äußert.

Wie das Beispiel zeigt, ist Verhalten daher *kanalisiert* in dem Sinne, daß es nicht durch die Unmenge potentieller Umweltinformation “abgelenkt”, aber trotzdem (evolutionär) erfolgreich ist.

Kanalisierung von Kognition

Da Verhalten durch Kognition erzeugt wird, können wir generell von einer Kanalisierung der Kognition sprechen. Sie ist bestimmten Restriktionen unterlegen und relativ unabhängig von äußeren Umständen. Für die Graugans existiert während des Rückholvorganges der Ethologie, der das Ei entfernt, einfach nicht.

Dasselbe gilt auch für menschliche Kognition. Psychologische Experimente zeigen, daß Dinge nicht außerhalb eines Kontexts existieren. Duncker (1935) verlangte von Versuchspersonen eine Kerze an einer Tür zu befestigen. Zur Verfügung standen Streichhölzer und eine Schachtel gefüllt mit Reißnägeln. Da die Versuchspersonen die Schachtel bloß als Behälter betrachteten, kam ihnen nicht in den Sinn, sie auszuleeren und mit Hilfe der Reißnägeln an die Tür zu heften, um so als Fundament für die Kerze zu dienen. Generell gesagt, scheint unser Denken über Dinge kanalisiert zu sein.

Aber auch prozedurale Kanalisierung ist feststellbar. Im Wasserumfüllproblem von Luchins (1942) waren Testpersonen gebeten, eine bestimmte Menge Wasser abzufüllen. Sie hatten 3 Krüge verschiedenen Inhaltes zur Verfügung, um durch Kombination von Addition und Subtraktion der Inhalte eine bestimmte Litermenge abzumessen. Einmal eine bestimmte Lösungsformel gefunden, die sich scheinbar für alle Aufgaben anwenden ließ, zeigten die Versuchspersonen keine Tendenz, einfachere Lösungen anzuwenden, wenn diese möglich waren. Ihr Denken war völlig in der gefundenen Strategie verhaftet. Luchins spricht von einer “Mechanisierung des Geistes”.

Die Kanalisierung der Kognition zieht sich selbst bis in den Bereich wissenschaftlichen Denkens. Ein Beispiel hierfür liefert die Entstehungsgeschichte der Keplerschen Gesetze. Trotz ihrer verhältnismäßigen Einfachheit und der bereits zu Keplers Zeiten ausreichend exakten Himmelsvermessung, benötigte Kepler 3 Jahrzehnte, um diese Gesetze aufzufinden (French 1999). Er war getrieben von der Überzeugung, geometrische Harmonien in den Umlaufbahnen der Planeten entdecken zu müssen, die ihn die tatsächliche Form von Planetenbahnen und das Verhältnis von Umlaufzeit und Sonnenabstand übersehen ließen. Sein Suchen nach den Harmonien war getragen von einem internen Gebäude von Vermutungen und Hypothesen, für die er keine Bestätigung fand. Er war erst dann erfolgreich, als er seine Vermutungen zu einem Satz von Hypothesen abänderte, die wir als die Keplerschen Gesetze kennen.

Wie die oben genannten Beispiele des Eirückholens und des Zurechtfindens in der eigenen Wohnung bei vollkommener Finsternis zeigen, weist die Kanalisierung der Kognition im generellen einen entscheidenden Vorteil auf: Sie führt zu einer Beschleunigung von Kognition, da sie die Anzahl von Entscheidungen einschränkt und trotzdem zu erfolgreichen Handlungen führt. Das steht im Kontrast zum Eimer-Paradigma der Informationsverarbeitung, wo Wichtiges aus einer immensen Datenflut ausgefiltert werden muß und dadurch einen Flaschenhals bei zeitkritischem Handeln darstellt.

Die Überlegenheit der Scheinwerfertheorie zeigt sich auch in den Arbeiten von Ernst von Glasersfeld (1987): Kognition besteht nicht im Abbilden der Realität, sondern in der Konstruktion viabler Hypothesen. Diese viablen Konstrukte sind dadurch gekennzeichnet, daß sie nur dann geändert werden müssen, wenn es der Kontext erzwingt. So ist seinem Beispiel folgend eine “Meerjungfrau” durchaus als zweibeiniges Wesen mit zusätzlichem Schwanz vorstellbar, solange sich dieses Konstrukt in den Kontext anderer Konstrukte widerspruchsfrei einreicht.

Oliver Sacks (1995) präsentiert uns das Beispiel des von Geburt an blinden Virgil, dem man im fortgeschrittenem Alter zum “Sehen” verhilft. Es stellt sich aber rasch heraus, daß

Virgils Blindheit nicht eine bloße Reduktion einer Sinneswahrnehmung darstellt, denn er hat größte Probleme, mit dieser für ihn neuen Form der Perzeption umzugehen. Blinde organisieren ihre Welt offensichtlich völlig anders. Sie erkennen Dinge durch Ertasten der Oberfläche in einer bestimmten Abfolge – ähnlich wie wir ein Musikstück nur sequentiell wahrnehmen können. Beginnend bei einem Startreiz werden gewisse Erwartungen produziert, die durch nachfolgende Tastempfindungen bestätigt werden oder das interne Aufrufen einer alternativen Wahrnehmungssequenz erfordern. Die Ähnlichkeit zu den weiter oben angeführten Beispielen des Zurechtfindens in der eigenen Wohnung, wie auch das ethologische Beispiel der Graugans ist klar ersichtlich und deutet auf einen gemeinsamen kognitiven Mechanismus hin. Wir leben in einer Welt der *Antizipation*, in der unser internes Gerüst an Vermutungen und Hypothesen uns zu Erwartungen führt. Wird eine Erwartung bestätigt, so können wir den eingeschlagenen “Denkweg” fortsetzen. Schlägt die Bestätigung fehl, so ändern wir entweder Abfolge der Vermutungen, oder modifizieren die Hypothesen selbst.

Jean Piaget (1954) und darauf aufbauend Ernst von Glasersfeld haben diesen Sachverhalt aus psychologischer Perspektive in eine “Schematheorie” gekleidet. Ein Schema erlaubt das Wiedererkennen einer bestimmten Situation, indem über gewisse Unterschiede zwischen gespeichertem Schema und aktuellem Kontext abstrahiert wird. Dieser Vorgang der Assimilation erlaubt also nur die Wahrnehmung dessen, was in bestehende Schemata paßt. Ein Schema assoziiert eine spezifische Aktivität mit der momentanen Situation, und führt daher zu der Erwartung, das gleiche Aktivitäten gleiche Resultate zeitigen.

Das kognitive System ist daher nicht durch den Sensorinput determiniert, sondern durch interne (unbeobachtbare) Erwartungen, welche die kognitive Aktivität kanalisieren. Dadurch, daß vertraute Situationen nur mehr erkannt, aber nicht immer wieder neue Reaktionen gelernt werden müssen, wird Kognition soweit beschleunigt, daß es in der Lage ist, in Echtzeit zu agieren.

Was kann die Artificial Life daraus lernen?

Die in dieser Arbeit behandelten Aspekte erläutern nicht nur kognitive Aktivität in bestehenden natürlichen Wesen, sondern weisen auch Implikationen für die Entwicklung künstlicher Geschöpfe auf. Ich fasse sie im folgenden zusammen und lege dar, auf welche Weise sie implementiert werden können.

Regeln als Bausteine der Kognition. Ethologische Beobachtungen zeigen, daß Verhalten und damit auch Kognition *regelhaft* sind. Sollen künstliche Systeme daher in Form von Regelsystemen realisiert werden? Schließlich ist evident, daß Regeln dem menschlichen Denken naturgemäß nahestehen: Nicht nur unsere Gesetze sind in Regeln abgefaßt, sondern auch das gesamte Spektrum alltäglicher Überlegungen (“Wenn ich dies tue, dann...”) bis hin zum Gebäude der Wissenschaft. Streng genommen ist selbstverständlich jedes Computerprogramm ein Regelsystem und das schließt auch neuronale Netzwerke ein, die oft als Alternative zu Regelsystemen genannt werden. Je expliziter jedoch die Regelstruktur einer Implementierung ist, desto besser vermag man die Dynamik eines solchen Systems zu verstehen, und sein Verhalten ist nicht in einer unerklärbaren Black-Box versteckt. Gerade die Artificial Life soll ein Werkzeug darstellen, das die Kluft zwischen Ethologen und Physiologen überbrückt: Zum Verständnis des internen Funktionierens einer komplexen nicht-trivialen Maschine ist die Kenntnis beider Funktionen, der Effektfunktion und der Zustandsfunktion nötig, andernfalls scheitert man beim Erklärungsversuch am Problem der Identifikation.

Man sollte nicht den Fehler begehen und ein Regelsystem generell mit einem symbolmanipulierenden System im Sinne der Artificial Intelligence (Newell und Simon 1976) gleichsetzen. Das Regelsystem in unserem Sinne ist nur das Vehikel für die Dynamik des

Systems, impliziert aber keinerlei Bedeutungszuweisungen. Das Symbol-Grounding-Problem wird damit nicht berührt.

Kanalisierung. Folgt man der Scheinwerfertheorie, so basiert ein kognitiver Prozeß auf dem Aneinanderreihen von Vermutungen und Bestätigungen innerhalb des internen Hypothesengerüsts. Ein Regelsystem eignet sich für die Implementierung, weil der Aktionsteil der einen Regel den Kontext für den Bedingungsteil der nachfolgenden Regel zur Verfügung stellen und damit Hypothesen aneinanderketten kann. Es sei hier an den Versuch zum Wasserumfüllproblem erinnert, bei dem das Erkennen des Kontextes (Assimilation) eine bestimmte Handlungssequenz aufruft. Während des Ablaufes einer Sequenz werden (wie bei der Graugans) all jene Umwelt Ereignisse vernachlässigt, die nicht zur Bestätigung der Richtigkeit der aktuellen Sequenz notwendig sind.

Semantik. Die im Prinzip der indifferenten neuronalen Kodierung festgehaltene Semantikfreiheit von Signalen kann in künstlichen Systemen durch eine Trennung zwischen kognitivem System und physiologischer Sinnesreizung realisiert werden. Der eigentliche kognitive Apparat darf nicht zwischen externer und interner "Information" unterscheiden.

Antizipation. Der aus der Scheinwerfertheorie abgeleitete Aspekt der erwartungsgetriebenen Kognition offeriert eine Alternative zur Bewertung eines kognitiven Systems. Viele Implementierungen stützen sich auf einen externen Fitneßbegriff, auf einen von außen vorgegebenen erstrebenswerten Zielzustand, mit dessen Hilfe der Lernvorgang gesteuert wird. Das Ziel des kognitiven Apparates besteht darin, den Abstand zwischen momentaner und zu erreichender Leistung zu minimieren. Die Vorgabe eines externen Fitneßkriteriums ist aber problematisch, wenn man jenen optimalen Zustand nicht kennt oder nicht angeben kann. Benützt man aber ein auf Antizipation beruhendes Kognitionssystem, so läßt sich der Erfolg direkt an den Bestätigungspunkten ablesen, sowie indirekt aufgrund der Echtzeitleistung bestimmen, die für kanalisierte Systeme höher ist.

Lernen. Die Schematheorie liefert Anregungen, wie der Lernvorgang in einem kognitiven System aussehen kann, unabhängig von externen Zielvorgaben. Werden Erwartungen in Bestätigungspunkten nicht erfüllt, so kann die betreffende Regel entweder vergessen werden, oder sie wird abgeändert, indem der Bedingungsteil der Regel ersetzt wird (sie somit in einem anderen Kontext gültig wird), oder die Komponenten des Aktionsteiles werden abgeändert.

Die Konstruktivistisch-Kognitive Architektur (Riegler 1994) stellt einen ersten Versuch der Implementierung eines künstlichen kognitiven Systems dar, welches einige der hier angeführten Aspekte und Richtlinien beinhaltet.

Zusammenfassung

Ich habe in diesem Artikel einige generelle Richtlinien für die Konstruktion von künstlichen Systemen aufgeführt, die der Fragestellung entwachsen, auf welche Weise ein Lebewesen mit der immensen Informationsflut in seiner Umgebung in Echtzeit umgehen kann. Die Erkenntnis spricht deutlich dafür, daß kognitives Handeln nicht in einem umfassenden Verarbeiten von Information besteht, sondern in der Errichtung eines internen Gebäudes von Vermutungen, die an gewissen Punkten auf ihre Angemessenheit überprüft werden. Diese Vorgehensweise ermöglicht auch kleinen Systemen eine ausreichend kognitive Schnelligkeit.⁵

5. Der in Fußnote 1 angestellte Vergleich zwischen menschlichen Schachspielern und Schachcomputer läßt erahnen, daß die "rechnerische Unterlegenheit" der menschlichen Spieler durch ähnliche Mechanismen (Heuristiken) wettgemacht wird, wie sie hier vorgestellt worden sind.

Das Vorhandensein von Bestätigungspunkten sagt natürlich nichts über den ontologischen Status einer Umwelt aus, da der kognitive Apparat aufgrund seiner operationalen Geschlossenheit keine primäre Unterscheidung zwischen Innen und Außen machen kann. Zugleich wird auch deutlich, daß die Frage nach dem Entstehen von Bedeutung in einem kognitiven System, natürlichen oder künstlichen Ursprungs, nicht von außen geklärt und umgekehrt ein Wesen nicht von außen in direkter Weise instruiert werden kann.

Literatur

- Bremermann, H. J. (1962) Optimization through evolution and recombination. In: Yovits, M. C. et al. (eds.) *Self-organizing systems*. Washington: Spartan Books.
- Clark, A. (1997) The Dynamical Challenge. *Cognitive Science* 4: 461–481.
- Dennett, D. C. (1984) Cognitive Wheels: The Frame Problem of AI. In: C. Hookway (ed.) *Minds, Machines, and Evolution: Philosophical Studies*. London: Cambridge University Press.
- Dreyfus, H. L. & Dreyfus, Stuart E. (1990) Making a Mind versus Modelling the Brain: Artificial Intelligence Back at a Branch-Point. In: Boden, M. (ed.) *The Philosophy of Artificial Intelligence*. Oxford Univ. Press.
- Duncker, K. (1935) *Zur Psychologie des produktiven Denkens*. Berlin: Springer.
- Foerster, H. von (1990) Kausalität, Unordnung, Selbstorganisation. In: Kratky, K. W. & Wallner, F. (eds.) *Grundprinzipien der Selbstorganisation*. Darmstadt: Wiss. Buchgesellschaft.
- French, R. (1999) When Coffee Cups Are Like Old Elephants, or Why Representation Modules Don't Make Sense. In: Riegler, A., Peschl, M. & Stein, A. von (eds.) *Understanding Representation in the Cognitive Sciences*. New York: Kluwer Academic / Plenum Publishers, pp. 93–100.
- Glaserfeld, E. von (1987) *Wissen, Sprache und Wirklichkeit*. Braunschweig: Vieweg.
- Harnad, S. (1990) The Symbol Grounding Problem. *Physica D* 42: 335–346.
- Lorenz, K. Z. & Tinbergen, N. (1939) Taxis und Instinkthandlung in der Eirollbewegung der Graugans. *Zeitschrift für Tierpsychologie* 2(1). Nachgedruckt in: Lorenz, K. (1984) *Über tierisches und menschliches Verhalten*. München: Piper.
- Luchins, A. S. (1942) Mechanization in Problem Solving. In: *Psychological Monographs* 54 (248).
- Maturana, H. R. (1982) *Erkennen: Die Organisation und Verkörperung von Wirklichkeit*. Braunschweig: Vieweg.
- Newell, A. & Simon, H.A. (1976) Computer Science as Empirical Enquiry: Symbols and Search. *Communications of the association for computing machinery* 19: 113–126.
- Peschl, M. & Riegler, A. (1999) Does Representation Need Reality? In: Riegler, A., Peschl, M. & Stein, A. von (eds.) *Understanding Representation in the Cognitive Sciences*. New York: Kluwer Academic / Plenum Publishers, pp. 9–17.
- Piaget, J. (1974) *Der Aufbau der Wirklichkeit beim Kinde*. Stuttgart: Klett.
- Popper, K. R. (1979) The Bucket and the Searchlight: Two Theories of Knowledge. In: *Objective Knowledge: An Evolutionary Approach* (rev. ed.). Oxford: Clarendon Press.
- Riegler, A. (1994) *Constructivist Artificial Life*. Dissertation an der Technischen Universität Wien.
- Sacks, O. (1995) *Eine Anthropologin auf dem Mars*. Reinbek: Rowohlt.
- Winograd, T. & Flores, F. (1986) *Understanding Computers and Cognition: A New Foundation for Design*. Norwood, NJ: Ablex.

Final Version of 11 Aug 1999