



Gotthard Günther [*]

Die "zweite" Maschine

Das erste "Werkzeug", das der Mensch in seiner Geschichte gebraucht haben mag, ist möglicherweise ein Stein gewesen, den er einem flüchtenden Tier nachwarf, um es zu töten. Die notwendigen Fähigkeiten, sich eines solchen Hilfsmittels zu bedienen, sind äußerst gering. Schon das höher entwickelte Tier hat sie. Ein Baboon ergreift eine Liane und schwingt sich mit ihrer Hilfe von Baum zu Baum. Ein Schimpanse weiß genau, dass man mit einem Stock einer Banane, die anderweitig nicht in Reichweite ist, habhaft werden kann. Und im Wald hinter dem Rhodes Memorial in Kapstadt bewerfen die Affen die dort im Auto heranfahrenden Besucher mit Steinen – in temperamentvoller Missbilligung moderner Autotechnik.

Aber weder Stock noch Liane oder Stein sind ein Werkzeug im menschlichen Sinn. "Nicht den zufälligen Gebrauch eines vor Augen liegenden Werkzeugs für nächste Zwecke, sondern die Herstellung eines Werkzeugs für einen fernen Zweck halten wir für menschliches Tun."^[1] Der Akzent liegt auf "Herstellung" und auf "fern". Das Tier stellt sein "Werkzeug" nicht mit einer fernen Zweckintention her. Es ergreift einfach, was vorhanden, d.h. vor seinen Händen, ist. Der Hebel aber, den der Mensch unter einen Steinblock schiebt, um diesen wegzurollen, ist etwas wesentlich anderes. Wesentlich ist jetzt nicht mehr der bloße Tatbestand einer Stange, die zu Händen ist, sondern die Tatsache, dass durch intelligente Placierung der Stange unter den Stein ein längerer und ein kürzerer Hebelarm hergestellt wird, deren Verhältnis erlaubt, einen längeren Arbeitsweg mit geringerem Kraftaufwand in einen kurzen Arbeitsweg mit höherer Kraftentfaltung zu verwandeln. Der Affe, der einen Stock oder eine Liane ergreift, verlängert damit nur künstlich seinen Arm – sein Verhältnis zur Umwelt ändert sich damit nicht im geringsten –, der Mensch aber, der einen Hebel benutzt, hat damit ein Stück seiner Intelligenz, resp. seines Bewusstseins, in die Außenwelt hineinprojiziert, wo es in dem Verhältnis von kürzerem und längerem Hebelarm für ihn arbeitet.

Das Tier, das einen Teil der Außenwelt werkzeughaft benutzt, verbessert damit nur den Funktionsbereich seines eigenen Körpers. Der Bussard, der thermische Aufwinde benutzt, um sich von ihnen rapide in die Höhe tragen zu lassen, und damit fliegerische Leistungen erzielt, zu denen er andernfalls außerstande wäre, projiziert damit keineswegs einen Teil seiner Handlungs- resp. Bewusstseinsintention in die Außenwelt und *verändert* sie dadurch. Alles freibewegliche Leben ist auf Erweiterung der eigenen Existenzkapazität angelegt, und das Mittel dazu ist die werkzeughafte Handhabung der Umwelt. Damit aber endet der Vergleich von Tier und Mensch in ihrem Gebrauch der Dingwelt unmittelbar. Kein Tier projiziert sich intelligenzmäßig über den Funktionscharakter des eigenen

* Erstveröffentlichung in: Edition und Nachwort *"Die zweite Maschine"*: Asimov, I., Ich, der Robot, Düsseldorf und Bad Salz 1952

Abdruck in: "Beiträge zu einer operationsfähigen Dialektik", Band 1, p.91-114, Felix Meiner Verlag, Hamburg, 1976.

¹ Arnold Gehlen, *Der Mensch*. Bonn, 1950, S. 68.

Körpers hinaus und in einen objektiven Sachzusammenhang hinein. Für tierische Existenz bleibt die Grenze zwischen Innenwelt (Mikrokosmos) und Außenwelt (Makrokosmos) unveränderlich. Nichts im Dasein des Tieres ist darauf angelegt, diese Grenze zu überschreiten.

Aus diesem Grunde können wir bestenfalls von einem werkzeughaften Gebrauch der Außenwelt durch den tierischen Organismus reden. Kein Tier aber wird ein selbständiges Werkzeug als unabhängige mittlere Existenzform zwischen sich und der Umwelt entwickeln. Eine solche Schöpfung erfordert, dass der Schöpfer etwas von sich abgibt und in einem objektiven Medium fixiert. Ein Obsidianmesser, das der Primitive vom Block abspaltet, ist als Material ein Stück des Naturzusammenhangs. Im Schlag, der es abspaltet, aber erhält es eine Form, die durch den Fernzweck seines Gebrauches bestimmt ist. Das heißt es hat nicht mehr natürliche, sondern künstliche Form, die einem bewussten Zweck entspricht. Ein solches Messer ist, wie Hegel sagen würde, ein Stück objektiver Geist. Objektiv insofern, als es ein materielles Stück der Außenwelt ist, und Geist insoweit, als die Natur von allein keine Werkzeuge hervorbringt und letzteres dem menschlichen Bewusstsein vorbehalten ist, das den bloßen Stoff durch bewusste Formung zur Stellvertretung seiner Handlungsintentionen zwingt. Damit kommt in die Idee der Existenz des Werkzeuges etwas Zwiespältiges. Das Werkzeug ist halb Natur und halb Geist. Es gehört voll weder auf die eine noch auf die andere Seite. Das Resultat der Zwiespältigkeit ist eine instabile Existenzform, die die Tendenz hat, sich von beiden Seiten abzulösen und etwas selbständiges Drittes zu bilden.

Dieser Ablösungsprozess vollzieht sich in der Geschichte der Technik, und diese ist bereits lang genug, um uns zu erlauben, aus ihr abzulesen, wie sich in ihr die Abtrennung des Werkzeugs von der Natur einerseits und vom Menschen andererseits vollzieht. Das ursprüngliche Material, das der Primitive für seinen Hebel verwendete, war ein natürlich gewachsenes Stück Holz. Die Pfeilspitze war ein Knochensplitter. Moderne Instrumente, die analogen Zwecken entsprechen, sind aus Stahl. Unsere Füllfederhalter sind aus Galalith, und in der Skifabrikation beginnen synthetische Kunststoffe die Naturhölzer zu verdrängen. Wen die Lust anwandelt, der kann sich heute ein Paar Skier aus Glasfiber kaufen. Die Natur, wenn sie sich selbst überlassen ist, bringt weder Stahl noch Galalith oder Glasfiber hervor. In der fortschreitenden Verwendung von Kunststoffen, die dem Verwendungszweck des Werkzeugs gemäßer sind als die Naturstoffe, findet eine ganz deutlich beobachtbare Ablösung des Werkzeugs aus dem Naturzusammenhang der Welt statt.

Derselbe Prozess der Ablösung vollzieht sich auf der subjektiven Seite. Das Werkzeug macht sich seinem Schöpfer gegenüber selbständig. Der Hebel, den der Primitive unter den Felsblock schob, bedurfte noch des Arms, der ihn niederdrückte. Das Mühlrad, das durch das Wasser des Dorfbachs betrieben wird, ist von der menschlichen Kraftquelle bereits unabhängig. Während nun die Ablösung des Werkzeugs von der Natur in der Schöpfung von Kunststoffen, d.h. einem ganz neuen Genus physischer Existenz, resultierte, führt auf der subjektiven Seite die Verselbständigung des Werkzeuges zur Maschine. Eine Wind- oder Wassermühle ist kein Werkzeug mehr, sondern eine Maschine. Eine Maschine ist nichts anderes als ein – innerhalb gewisser Grenzen – autonom gewordenes Werkzeug.

Unsere Sprache drückt den Verselbständigungsprozess, der sich in dieser Entwicklung vollzieht, sehr präzise aus: ein Werkzeug wird *gehandhabt*. Eine Maschine wird *bedient*. Das gilt wenigstens von den älteren Maschinentypen. Ein neuerer Maschinentypus, wie der Servomechanismus, erfordert nicht einmal mehr Bedienung, sondern nur noch gelegentliche Wartung. Ein Thermostat, der die Temperatur in einer neuzeitlichen Wohnung regelt, braucht nicht mehr bedient zu werden, um sinngemäß zu funktionieren. Dieser Mechanismus dirigiert seinen Arbeitsmodus selbst.

Die technische Entwicklung geht also vom nicht-automatischen Werkzeug oder Elementarmechanismus (Töpferscheibe, Spinnrad usw.) zur halbautomatischen Maschine und von da zum vollautomatischen maschinellen Arbeitsaggregat. Ein Automobil ist z.B. ein halbautomatischer Mechanismus. Nur noch einige wenige seiner Arbeitsvollzüge werden von dem bedienenden Fahrer reguliert. Andere, wie die Verstellung des Zündzeitpunktes, Ventiltätigkeit, Generatorbetrieb, Kuppeln und Gangwechsel (in Modellen mit vollmechanisierter Kraftübertragung) besorgt die Maschine von allein. Die thermostatisch geregelte Öl- oder Stromheizung hingegen ist voll-automatisch. Ihr Arbeitszyklus ist vom Menschen unabhängig. Ihr Schöpfer hat nur noch damit zu tun, *dass* sie für ihn arbeitet. In der Ausführung der Arbeit dirigiert sich die Maschine selbst. Vollautomatische Maschinen werden heute in der Industrie weitgehend verwendet. Einige amerikanische Lebensmittel-firmen weisen in ihrer Reklame mit Stolz darauf hin, dass ihre Esswaren nie von menschlichen Händen berührt worden sind. Die erste Hand, die mit der Nahrung in direkten Kontakt kommt, ist die des Käufers, der das sauberlich verpackte Paket öffnet.

Voll-automatische Maschinen, wie z.B. bestimmte amerikanische Verkehrssignalanlagen, die ihren Lichtwechsel danach richten, in welcher Straße die größere Anzahl von Fahrzeugen auf Durchfahrt wartet, werden gelegentlich "Robots" genannt. Der Name ist in Europa zuerst etwa 1923 aufgetaucht und geht auf einen Roman des tschechischen Schriftstellers Karel Capek zurück, in dem solche selbsttätigen Maschinentypen eine Rolle spielen. Die Etymologie des Wortes ist tschechisch (*robit*). Der Robot ist der mechanische Arbeiter.

Von Europa aus ist der Ausdruck in den amerikanischen Sprachschatz übernommen worden. Hier aber hat sich ein sehr tiefgehender Wandel in der Bedeutung dieses Wortes vollzogen. Ein Robot ist, wie Asimovs Buch sehr deutlich zeigt, viel mehr als nur "eine vollautomatische Maschine, die dem Menschen seine Arbeit abnimmt". Der amerikanische Sinn des Wortes "Robot" weist auf einen neuen, bisher nicht dagewesenen Maschinentyp hin, dessen generelle Konstruktionsidee sich radikal von den aus dem elementaren Werkzeug entwickelten Apparaturen, die wir heute Maschinen nennen, unterscheidet. Wir wollen den bisherigen Maschinentyp, da sein Arbeitsmodus von dem archimedischen Hebelprinzip ausgeht, die archimedisch-klassische Maschine nennen.

Das handgreifliche Kennzeichen dieses Maschinentyps ist, dass er mechanisch bewegliche Teile (Hebel, Achse, Rad, Schraube) besitzt, und dass er seine Arbeit durch die Bewegung dieser Teile verrichtet. In diesem Sinn besteht zwischen einem Rolls-Royce und einem Baumstamm, den der Naturmensch auf den Boden legt, um vermittels seiner eine Last zu rollen, nicht der geringste Unterschied.

In unserer Gegenwart aber beginnen die Anfänge eines neuen Maschinentyps aufzutreten, einer Arbeit leistenden Apparatur, die keine mechanisch beweglichen Teile mehr hat und deshalb auch nichts durch Bewegung solcher Teile verrichtet. Ein erstes (äußerst primitives) Beispiel dieses Typs ist der elektrische Transformator, der wesentlich aus nichts anderem als zwei Drahtwicklungen und einem Stück Eisen besteht. Dieser Mechanismus hat keine sich bewegenden Teile mehr. Wenigstens nicht in dem üblichen Sinn des Wortes. Alle arbeitsleistende Bewegung erfolgt hier durch Atome bzw. Elektronen und magnetische Felder. Wir wollen diesen Maschinentypus, weil er nicht mehr auf dem archimedischen Hebelprinzip beruht, die trans-klassische oder nicht-archimedische Maschine nennen. Die sich in einem solchen "Mechanismus" vollziehenden Arbeitsvorgänge vollziehen sich nämlich nicht mehr nach klassisch-mechanischen Prinzipien.

Der Prototyp der archimedischen Maschine ist der menschliche Körper *mit seinen beweglichen Gliedmaßen*. Der Mensch, der eine archimedische Maschine baut, wiederholt die Arbeitsschematik seines eigenen Körpers in einem zweiten, künstlichen "Körper" in der Außenwelt. Die Analogie des arbeitenden Armes ist z.B. in einem Explosionsmotor in dem Bewegungsrhythmus von Kolben, Pleuelstange und Kurbelwelle einfach nicht zu verkennen. Dieser Parallelismus zu arbeitenden Körpergliedern aber versagt vollkommen, wenn wir uns die Arbeitsweise eines elektrischen Transformators oder des vor einigen Jahren von den Bell-Laboratorien entwickelten Transistors vorstellen wollen.

Nun besitzt der Mensch aber ein Organ, das nicht klassisch-mechanisch mit beweglichen Teilen (wie z.B. das Herz), sondern in einer sehr andersartigen Weise, nach trans-klassischen Prinzipien arbeitet. Es ist dies das Gehirn. Die Arbeitsweise der klassischen Maschine folgt dem Vorbild des arbeitenden Armes (samt Hand). *Die Idee der trans-klassischen Maschine aber erwächst aus den technischen Forderungen, einen Mechanismus zu entwickeln, der nach der Analogie des menschlichen Gehirns arbeitet.* Eine solche erfolgreich durchgeführte Konstruktion wäre ein Robot im neuen amerikanischen Sinn des Wortes. Damit aber ist ein neuer, in der bisherigen technischen Geschichte des Menschen nicht dagewesener Maschinentyp konzipiert. Die allgemeine Theorie eines solchen nicht-archimedischen Mechanismus wird in einer neuen, kürzlich in den Vereinigten Staaten entwickelten wissenschaftlichen Disziplin, der sogenannten Kybernetik« untersucht. Der Name dieser neuen Wissenschaft ist aus dem Griechischen entlehnt und von ὁ κυβερνήτης, der Steuermann abgeleitet. Eine kybernetische oder nicht-archimedische Maschine ist nämlich eine solche, die nicht mehr physische Arbeitsvorgänge produziert (das bleibt weiterhin dem klassischen Mechanismus überlassen), sondern die solche Arbeitsvorgänge dirigiert und "kritisch" steuert. Maschinen, die diese Idee auch nur annähernd verwirklichen, existieren heute noch nicht, aber ihre theoretische Möglichkeit ist anerkannt. Man arbeitet an ihrer Konstruktion. Um die Grundkonzeption einer solchen Maschinenidee zu verstehen, müssen wir uns vorerst fragen: was liefert eine solche Maschine? Der archimedische Typ des Mechanismus liefert physische Arbeitsvorgänge. Er dirigiert (steuert) aber diese Arbeitsvorgänge nicht auf sinnvolle Zwecke hin. Das bleibt uns überlassen. Das neue maschinen-theoretische Problem ist also das folgende: welches Produkt kann durch eine nicht-archimedische Maschine geliefert werden, das seinerseits imstande ist, die Arbeitsvorgänge der klassischen Maschine sinngemäß zu lenken?

Eine kurze Überlegung wird uns zeigen, dass es darauf nur eine einzige Antwort gibt, die ganz einfach und von zwingender Evidenz ist. Wie werden archimedische Maschinen vorläufig gesteuert, da kybernetische Mechanismen vorerst noch nicht existieren? Das Auto ist eine archimedische Maschine, und sein praktischer Gebrauch ist – zumindest theoretisch – jedermann geläufig. Ein Auto in Betrieb ist ein Auto im Verkehr und seine sinnvolle Arbeitsweise wird uns dadurch garantiert, dass ein Fahrer am Steuer sitzt, der Verkehrsregeln wie:

Linksfahren verboten!
Rotes Licht: Halt!
Grünes Licht: Fahren!
Höchstgeschwindigkeit: 30 km/h!
Fahrzeug von rechts hat Vorfahrt!

durch entsprechende Handlungen seinerseits seiner Maschine mitteilt und dadurch die Arbeitsvorgänge in ihr modifiziert. Kurz, der Fahrer lenkt sein Auto durch den Verkehr. Dies geschieht auf Grund von Verkehrsregeln. Verkehrsregeln aber sind Informationsdaten, die der Fahrer als allgemeine Richtlinien der Verkehrsordnung übernimmt und dann im Einzelfall der konkreten Fahrtsituation als detaillierte Information selbst produziert. Die abstrakte Verkehrsvorschrift mag zwar sagen: Grünes Licht Fahren! Die Kreuzung aber steht noch voll Wagen, die den Weg bisher nicht freigegeben haben. Das ist modifizierendes Informationsmaterial, das das Bewusstsein des Fahrers verarbeiten muss und das schließlich in einer "kybernetischen" Bewegung endet, nämlich mit einem Druck auf die Fußbremse.

Dieses Beispiel zeigt deutlich, dass ein archimedischer Mechanismus auf Grund von Information gesteuert wird. Teilweise kommt diese Information direkt von der Maschine her. Die Instrumente am Armaturenbrett: Tachometer, Öldruckmesser, Amperemeter usw. liefern nichts weiter als Informationsdaten, und dieses informative Datenmaterial ist für den Fahrer notwendig, um die Maschine sinngemäß zu bedienen. Die Idee der kybernetischen Maschine zielt also auf die konstruktive Verwirklichung eines Mechanismus, der Daten aus der Außenwelt aufnimmt, sie als Information verarbeitet und dieselbe in Steuerungsimpulsen dann an die klassische Maschine weitergibt. Das letztere ist übrigens nicht durchaus notwendig. Eine Rechenmaschine z.B. liefert unmittelbar arithmetische Information.

Wir können also ganz allgemein im Rahmen einer Philosophie der Technik sagen: der Mensch hat bisher in seiner technischen Entwicklung zwei grundverschiedene Ideen der Maschine konzipiert. Die erste ist die klassisch-archimedische Maschine, deren Zweck ist, Arbeit zu produzieren. Neben diese ist die Idee der "zweiten" Maschine getreten, von der man nicht mehr Arbeit, sondern Information erwartet. Die "erste" Maschine ist in Analogie zum menschlichen Arm (und Hand) entworfen worden, von der zweiten wird erwartet, dass sie eine technische Reproduktion des menschlichen Gehirns darstellen soll. Denn nur das Gehirn verarbeitet Information. Die "erste" Maschine ist heute eine historische Realität, die "zweite" vorläufig nur ein technisches Ideal, dessen progressive

Realisierung^[2] noch in sehr weiter Ferne steht. Transformatoren, Transistoren und ähnliche auf elektro-magnetischen Prinzipien beruhende Maschinen stellen noch nicht einmal eine erste Annäherung an das neue Ziel dar. Sie sind technische Zwischengebilde, die sich dadurch von dem archimedischen Maschinentyp unterscheiden, dass in ihnen der Mechanismus in subatomare Bereiche verlegt worden ist. Insofern folgen sie in der Tat einem nicht-klassischen Prinzip. Ihre "mechanische" Arbeitsweise ist bereits die der projektierten kybernetischen Maschinen. Sie liefern aber keine Information. Insofern sind sie klassisch. Auch die heutigen Rechenmaschinen bis hinauf zum kompliziertesten Differentialanalysator sind noch keine echten trans-klassischen Konstruktionen. Sie *transformieren* nur gelieferte Informationen, aber sie produzieren keine.

In unserem Beispiel von dem Autofahrer im Stadtverkehr aber wird durch das Gehirn Information *produziert*. Die durch die Verkehrsordnung gelieferte und auswendig gelernte Information reicht nicht im entferntesten aus, um den Wagen, im Sinn dieser selben Verkehrsordnung sach- und sinngemäß zu bedienen. Wir können – um den entscheidenden Tatbestand, auf den es hier ankommt, klarer herauszustellen – einmal etwas übertreiben und sagen: die Verkehrsordnung liefert überhaupt keine Information, sondern nur Richtlinien, gemäß denen die für die Steuerung des Wagens unbedingt notwendige Verkehrsinformation produziert werden soll. Selbstproduktion von Information, gleichgültig ob sie partiell oder total ist, aber setzt Bewusstsein voraus.

Damit konzentriert sich das Problem der nicht-archimedischen Maschine auf das des "mechanical brain"^[3], d.h. des mechanischen Bewusstseins. Und es erhebt sich die den heutigen Durchschnittsmenschen geradezu ungeheuerlich anmutende Frage: Lässt sich Bewusstsein trans-klassisch mechanisch konstruieren? Die erste Reaktion auf diese Frage wird für jeden klassisch erzogenen und im Bereich einer ontologischen Metaphysik aufgewachsenen Zeitgenossen ein ganz emphatisches "Nein" sein. Tatsächlich aber ist die Frage für ein unbefangenes, vorurteilsloses Denken vollkommen offen. Was allein unwiderleglich feststeht, ist, dass es nicht möglich ist und nie möglich sein wird, ein volles menschliches *Selbstbewusstsein* als Robotgehirn zu entwerfen. Und zwar aus dem folgenden Grunde: die Logik bzw. Mathematik, in der ein solcher "mechanical brain" beschrieben wird, muss von einem höheren Sprachtypus^[4] sein als derjenige, den das Robotgehirn braucht, um seine Begriffe zu produzieren. In der Ausdrucksweise der symbolischen Logik: die Konstruktion eines Robots muss in einer Sprache erfolgen, die relativ zu der Sprache, in der ein Robot "denkt", eine Metasprache ist. Nun gibt es aber zu einer Sprache, die Begriffe wie "Ich", "Du" oder "Selbst"

² Es muß darauf aufmerksam gemacht werden, dass eine solche Realisierung immer nur partiell erfolgen kann. Einen kybernetischen Mechanismus, der das Gehirn *vollendet* reproduziert, wird man nie konstruieren.

³ Wir benutzen für "mechanical brain" den männlichen Artikel, weil "brain" mit dem deutschen Ausdruck "der Brägen" für Gehirn etymologisch ist. "Brägen" ist provinziell und wird vornehmlich im Fleischergewerbe gebraucht.

⁴ "Sprachen" in dem hier intendierten Sinn sind nicht nur Deutsch, Englisch, Chinesisch, usw., sondern ebenfalls Mathematik, Logik, kurz alle Ausdruckssysteme, in denen sinnvolle Mitteilungen gemacht werden können.

als logisch relevante Ausdrücke enthält, keine Metasprache mehr. Eine solche Sprache ist von höchstmöglicher logischer Ordnung. Wenn also ein Konstrukteur versuchte, einem "mechanical brain" die eben genannten Begriffe und damit ein Denken in einer Sprache höchstmöglicher Ordnung einzubilden, dann bliebe ihm keine Metasprache mehr, in der er ein solches Robotgehirn entwerfen könnte. Umgekehrt: reserviert der Konstrukteur einen solchen Sprachtypus für die *Darstellung* seines Entwurfes, dann kann er dem Entwurf selber nur ein niederes Sprachniveau, in dem solche Worte (Begriffe) noch nicht auftreten, zuschreiben.^[5] Ein Gehirn aber, das den Begriff "Selbstbewusstsein" prinzipiell nicht konzipieren und in seiner Sprache bilden kann, hat auch kein Selbstbewusstsein.

Es wird also nie möglich sein, einen Robot, der Selbstbewusstsein besitzt, zu konstruieren, weil ein "mechanical brain", der Worte wie "Ich" und "Selbst" gebrauchen kann und weiß, was sie bedeuten, eine Sprache spricht, zu der es keine Metaphysik mehr gibt, in der sein technischer Entwurf vom Konstrukteur konzipiert werden könnte. Wenn beide die gleiche Sprache sprächen, dann wären Schöpfer und Geschöpf einander geistig ebenbürtig. Dies ist absurd.

Wie steht es aber mit der technischen Konstruktion von einfachem Bewusstsein? Das bisher Gesagte bezieht sich ausschließlich auf das Selbstbewusstsein, d.h. auf doppelreflektierte Erlebnisphänomene. In der Tat ist die Frage heute noch völlig offen, ob sich einfaches Bewusstsein, also ohne weitere Rückreflexion auf das erlebende Ich, jemals mit technischen Mitteln herstellen ließe. Dass es partielle Bewusstseinszustände ohne Ich-Erlebnisse gibt, wird nirgends ernsthaft bestritten. Wer jemals aus einer Chloroformnarkose aufgewacht ist, kennt jenen Übergangszustand, in dem das Bewusstsein dadurch zum Leben erwacht, dass es beginnt, sich wieder mit Inhalten zu füllen. Der krönende Inhalt aber, die Rückbeziehung auf die eigene Bewusstseinstätigkeit, d.h. des Ich-Erlebnisses, fehlt und will sich nicht erzwingen lassen. Der inverse Vorgang, freilich nicht so markant und eindrucklich wie in den künstlichen Bewusstseinsveränderungen durch die Narkose, lässt sich bei dem natürlichen Prozess des Einschlafens

⁵ Für denjenigen Leser, der an diesem Grundproblem einer allgemeinen Theorie eines "mechanical brain" näher interessiert ist, seien im folgenden die vier logisch möglichen Sprach- und Ausdruckssysteme mitgeteilt. Man unterscheidet 1. Sprachen, in denen alle Ausdrucksvariablen zu einer und derselben semantischen Kategorie gehören; 2. Sprachen, in denen die Anzahl der die Variablen umfassenden Kategorien größer als 1, aber stets endlich ist; 3. Sprachen, in denen die Variablen zu unendlich vielen semantischen Kategorien gehören, wobei aber die Ordnung dieser Variablen eine im vornhinein gegebene natürliche Zahl nicht überschreitet, und schließlich 4. Sprachen, die Variable beliebig hoher Ordnung enthalten. (Vgl. Alfred Tarski, *Der Wahrheitsbegriff in den formalisierten Sprachen*. *Studia Philosophica*, Leopoli 1935, bes. S. 81). Alle höher entwickelten Umgangssprachen, die Begriffe wie "Ich" und "Selbst" enthalten, gehören der vierten Sprachordnung an. Jede dieser Sprachordnungen ist relativ zu den niederen eine Metasprache. Und man kann über eine Sprache nur in einer ihr übergeordneten Metasprache wissenschaftlich exakt sprechen. Auf dem Niveau des vierten Sprachtypus, der auch Universalsprache genannt wird, kann man über alles sprechen; allerdings mit der höchst beträchtlichen Einschränkung, dass man in der Logik, die diese Sprache beschreibt, Paradoxien und Antinomien in Kauf nehmen muss, wenn man über "Tatbestände" spricht, deren Begriffe erst auf diesem Sprachniveau sich bilden. Solche Begriffe sind z.B. "Ich" oder "Selbstbewusstsein". Paradoxien aber sind nicht als technische Objekte konstruierbar.

beobachten. Das Ich-Erlebnis verschwindet zuerst. Die anderen Bewusstseinsinhalte später. Ein Bewusstsein, das noch Inhalt hat, *ist bewusst*. Das heißt, wir haben in diesen Übergangszuständen Bewusstseinszustände ohne ein konkomitierendes Ich.

Ebenso dürfte es richtig sein zu sagen, dass das Kind in seinem frühesten Entwicklungsstadium unmittelbar nach der Geburt zwar Bewusstsein, aber kein Selbstbewusstsein hat. Das Tier schließlich bleibt auf der Stufe des einfachen, nach außen gerichteten Bewusstseins überhaupt stehen und erreicht (vermutlich) niemals den selbstbewussten Erlebnisraum. Soweit lässt sich sagen, dass Bewusstsein sehr wohl ohne Selbstbewusstsein "psychisch" realisiert werden kann. Warum also nicht in einem "mechanical brain"?

An dieser Stelle macht unser im seelischen Determinationsraum der klassischen Ontologie erzogenes Weltgefühl einen anderen und tieferen Einwand: zugegeben, dass Bewusstseinszustände ohne Selbstbewusstsein realisiert werden können. Aber jedes Bewusstsein muss, selbst wenn es nirgends an die Oberfläche tritt, ein Subjekt seiner Akte, eine "Seele" besitzen. Im Tier oder im Kind der ersten Lebensjahre liegt sie nur unter der Erlebnisschwelle, aber sie ist potentialiter immer da. Bewusstsein ist nicht – wie der vulgäre Materialismus annimmt – eine Funktion eines physischen Agens, also ein Endprodukt aus Stoff und Energie, sondern vielmehr die Eigenschaft eines unbekanntes X, das wir in theologischer Terminologie "Seele" nennen.^[6]

Diese theologische Auffassung unterschreiben wir hier völlig. In ihr formuliert sich ein existentieller oder metaphysischer Sachverhalt, der kaum ernsthaft bestreitbar ist. Die theologische Formulierung jedoch, so wahr sie auch essentiell sein mag, ist äußerst vage. Sie lässt die faktische Relation, die die "Seele" zum Bewusstsein hat, ganz offen und macht sie den verschiedenen Interpretationen zugänglich. Ein Symptom dafür ist die schwankende Stellung, die das theologische Denken der Frage gegenüber einnimmt, ob Tiere eine unsterbliche Seele haben. Der Grund für die bestehende Unsicherheit in der Beantwortung dieses sehr heiklen Punktes ist die durchaus richtige Einsicht, dass, wenn Tiere kein Selbstbewusstsein haben, die Relation ihrer hypothetischen "Seele" zu ihrem Bewusstsein eine ganz andere sein muss als beim Menschen.

Der Leser sei hier an den Anfang des amüsanten Romans von Anatole France "Die Insel der Pinguine" erinnert. Der sich im Boote der Insel nähernde St... hält die am Ufer sitzenden Pinguine irrigerweise für die heidnischen Bewohner des Landes und tauft sie summarisch. Dadurch erhalten die Pinguine (christliche) Seelen, und im Jenseits entsteht jetzt das sehr ernste Problem, ob die Pinguine zum Eintritt in den Himmel berechtigt sind oder nicht. "Seele" ist hier offenbar etwas, das im Taufakt "mit-geteilt" werden kann. Wären es statt der Pinguine menschliche Wesen gewesen, dann hätte sich im Taufakt das mystische Verhältnis zwischen

⁶ Norbert Wiener formuliert diesen Sachverhalt in seinem epochemachenden Buch "Cybernetics" (New York, S. 155) in der folgenden für einen Ingenieur typischen Weise: Information is information, not matter or energy." (Information - d.h. Bewusstseinsinhalte - ist Information und nicht Materie oder Energie). Information oder auch Sinn ist eben die empirische Weise, in der sich uns jenes geheimnisvolle X offenbart.

Taufendem und Täufling in anderer Weise arrangiert. In diesem Fall wäre die Seele nicht auf das getaufte Individuum übertragen worden, sondern die im Individuum bereits wohnende Seele wäre nur christianisiert worden. Die stillschweigende Voraussetzung dieser Episode ist: Menschen haben *Selbstbewusstsein*, d.h. ihre Seele wohnt bereits in ihnen, Tiere haben keins, also ist ihr Bewusstsein, was Information anbetrifft, ferngesteuert. Der Steuerungsimpuls ihres Bewusstseins kommt aus der Umwelt. Umwelt aber ist relativ zum Individuum bzw. seinem Bewusstsein Transzendenz.

Das Tier hat nach dieser Auffassung ebenso eine transzendente ."Seele" wie der Mensch; nur muss "Transzendenz" in diesem Fall anders definiert werden. Die menschliche Seele ruht unerreichbar im "Innern"; sie ist, nach einem adäquaten Sprachgebrauch, introszent. Die tierische Seele ist extroszent. Sie liegt im "Außen". In beiden Fällen liegt das, was wir als Subjekt des Bewusstseins zu hypostasieren gezwungen sind, nicht auf der Existenz- und Aktionsebene des Bewusstseins selbst.

Solange diese Auffassung von den verschiedenen Transzendenzmöglichkeiten des Subjekts gegenüber seinem Bewusstsein nicht zwingend widerlegt ist (und eine solche Widerlegung existiert bis heute nirgends), kann die theoretische Möglichkeit nicht von der Hand gewiesen werden, dass es prinzipiell möglich ist, Bewusstsein in einem "mechanical brain« durch konstruktive Methoden hervorzurufen. Ein solches Bewusstsein eines Robotergehirns würde auch eine "Transzendenz" besitzen. Aber diese Transzendenz wäre nicht die Introszendenz des menschlichen Bewusstseins (der Robot hat kein Ich im "Innern") sondern eine zweite Form Extroszendenz wie beim tierischen Bewusstsein. Nur ist im Tier die Umwelt das Lokale dieser Transzendenz, während für den Robot dieselbe – im Bewusstsein seines Konstrukteurs liegt. Der vollendete Robot hätte ein "Ich". Dasselbe aber wäre zurückverlegt in das stellvertretende Ich seines Schöpfers!

Es braucht wohl kaum ausdrücklich darauf hingewiesen werden, dass mit den heute zur Verfügung stehenden technischen Mitteln solche Ziele auch nicht im entferntesten verwirklicht werden können. Was in hundert Jahren technisch möglich sein wird, kann momentan niemand sagen. Die heute gebauten Modelle logischer und mathematischer Komputiermaschinen wie ENIAC, EDVAC, UNIVAC und andere jagen sich in einem solch rasenden Tempo, dass, wie ein auf diesem Forschungsgebiet tätiger amerikanischer Gelehrter kürzlich bemerkte, eine derartige Maschine, sobald sie fertiggebaut ist, auch schon als veraltet angesehen werden muss. Es lässt sich deshalb in diesem Kommentar vom technischen Standpunkt aus kaum etwas sagen, was nicht beim Druck dieser Zeilen längst überholt sein wird.

Anders aber liegen die Dinge, soweit die allgemeine logische Theorie dieser Maschinen in Betracht kommt. Hier lassen sich Feststellungen machen, die allgemein und unabhängig von den technischen Mitteln, sie zu realisieren, richtig sind und deshalb nicht überholt werden können. Eine solche Feststellung ist, dass keinerlei logisch-theoretische Gründe dagegen sprechen, dass Bewusstsein (*nicht Selbstbewusstsein*) technisch konstruierbar ist, wenn die Bewusstseinsprozesse eines solchen Mechanismus Fernsteuerung implizieren. Die logischen Gründe für eine solche Behauptung fußen auf der Tatsache, dass sich eine wissenschaftlich

zugreifende und genügend exakte Definition von "Bewusstsein" auf der Ebene der dritten Sprachordnung geben lässt. Der Konstrukteur hat dann für seine Arbeit den vierten und letzten Sprachtypus als seine "Metasprache" zur Verfügung. Damit wird die Sprache dritter Ordnung relativ zur vierten zu einer reinen Objektsprache. Das bedeutet, dass man bei der Definition von "Bewusstsein" nur Objektbegriffe wie "Materie", "Energie" und "Information" zu verwenden braucht und dass Subjektbegriffe^[7] wie "Erlebnis", "Ich", "Du", "Seele" usw., überflüssig geworden sind.

Mit anderen Worten: es bestehen keine prinzipiellen Schwierigkeiten, die Produktion von Bewusstsein als einen rein "mechanischen" Prozess eines entsprechenden Instrumentes – des Gehirns oder des "mechanical brain" – aufzufassen. Damit aber ist wenigstens die theoretische Möglichkeit der Konstruktion einer solchen, uns heute völlig phantastisch anmutenden Maschine gegeben. Wohlgermerkt, ein solcher Apparat würde alle Bedingungen für das Zustandekommen von Bewusstsein erfüllen – nicht aber diejenigen für Selbstbewusstsein. Sein Bewusstsein könnte also nur in Analogie zu dem eines Tieres oder Kindes im frühesten Lebensstadium gesetzt werden. Andererseits aber würde das Bewusstsein eines "mechanical brain" von dem eines Tieres oder Kindes sich insofern unterscheiden, als es trotz seiner dumpfen Bewusstseinsstufe ohne weiteres schwierigen Problemen der Differential- und Integralrechnung gewachsen wäre.

Es bleibt in dieser Darstellung des kybernetischen Robotproblems nur noch übrig, kurz zu bemerken, wie weit die Technik der modernen Kalkulatoren auf diesem Weg bereits vorgeschritten ist. Hier ist folgendes festzustellen: alle bisher konstruierten Maschinen folgen einer Logik, die dem allerersten Sprachtypus entnommen ist. Dies ist die Logik, die den Relationen toter Objekte zueinander entnommen ist. Folglich sind diese Maschinen auch so tot wie ein Stein. Es sind aber in diese Maschinen bereits einige technische Details hineinkonstruiert worden, die der nächst höheren Sprach- und Logikstufe angehören. Das heißt, die modernsten Maschinen besitzen bereits die mechanischen Vorbedingungen für Gedächtnis (und diese Anlage teilt sich ganz wie in der menschlichen Psyche in ein permanentes und ein temporäres Gedächtnis), für Lernfähigkeit und für "Gestalt"wahrnehmung.^[8] Sie besitzen ferner in beschränktem Maß Entscheidungsfähigkeit (Auswahl arithmetischer Routinen) und können komplizierten Instruktionen, die ihnen in Maschinensprache« gegeben werden, folgen. Schließlich besitzen sie die bemerkenswerte mechanische Fähigkeit, abstrakte Begriffe in sinnvolle Handlungsrou-tinen umzusetzen. Zu diesen allgemeinen Funktionen, die jedes Bewusstsein kennzeichnen, treten sehr spezifische logische Fähigkeiten, deren Vollzug mechanisch reproduzierbar ist. Eine von Theodore Kalin und William Burkhart gebaute Maschine, die nicht größer als ein Fernsehapparat ist, prüft die Richtigkeit von logi-

⁷ Das heißt Begriffe, die auf ein hypothetisches Subjekt hinweisen und dieses für ihre Erklärung benötigen.

⁸ Dies ist eine Fähigkeit, die wir z. B. brauchen, wenn wir die schwer leserliche Handschrift eines Briefes entziffern wollen. Wir bemühen uns dann, in dem willkürlichen Duktus der Handschrift die Idealgestalten der Buchstaben wiederzuerkennen. Das erste Modell einer Maschine, die die Fähigkeit der Gestaltwahrnehmung hat, ist kürzlich von David Shepard (Falls Church, Va.) konstruiert worden.

schen Schlüssen, die sich aus Aussagengruppen ergeben. Ein weiteres Modell befindet sich in Bau, das Probleme aus Spezialgebieten des logischen Funktionenkalküls lösen kann und das in beschränktem Maß zu Existenzurteilen⁹] fähig ist.

Allgemein wird auch von konservativen Kybernetikern zugegeben, dass es theoretisch möglich ist, jede Bewusstseinsfunktion mechanisch zu wiederholen. Denn wenn die Funktionsweise *eines* Bewusstseins-elementes mechanisch interpretierbar ist, dann muss das Gleiche auch von allen anderen gelten. Nur eine Fähigkeit, so wird einstimmig betont, wird man niemals konstruktiv duplizieren können! Es ist die schöpferische Tätigkeit des menschlichen Bewusstseins. Wozu wir hier bemerken wollen, dass spirituelle Produktivität vermutlich überhaupt keine Bewusstseinsfunktion ist und ausschließlich dem Bereich des ohnehin nicht reproduzierbaren Selbstbewusstseins angehört.

Abgesehen davon aber ist eine kybernetische Entdeckung gemacht worden, die den projizierten "mechanical brains" einen sehr "menschlichen" Anstrich gibt. Dr. John R. Pierce, ein in den Bell Laboratorien arbeitender Forscher, hat kürzlich nachgewiesen, dass jedes denkende Bewusstsein ein Zufallselement¹⁰] enthält. Es wird dann mit außerordentlichem Scharfsinn gezeigt, dass zu erfolgreichem Denken gehört, dass ein Bewusstsein (oder "mechanical brain") langsam lernen muss und dass es weiterhin die Fähigkeit besitzen muss, früher Gelerntes entweder zu ignorieren oder zu vergessen. Darüber hinaus hat Pierce eine Konstruktion entworfen, die das Problem praktisch löst und zeigt, wie man in ein Robotgehirn diese Eigenschaften einbauen kann. Die praktische Durchführung ist übrigens so einfach, dass jeder halbwegs begabte Elektrotechniker die dazugehörigen Diagramme lesen kann. —

Warum nun, wird der Laie fragen, kann man heute noch kein Bewusstsein konstruieren, wenn einzelne Bewusstseinsfunktionen bereits mechanisch reduplizierbar sind? Die Antwort ist: man würde auch selbst dann noch kein solches Bewusstsein konstruiert haben, wenn man *alle* überhaupt existierenden Funktionen eines lebendigen Bewusstseins in tadellos arbeitenden Mechanismen wiederholt hätte. Das hat Plato schon gewusst, wenn er im Dialog Theätet darauf hinweist, dass dadurch, dass im Trojanischen Pferd in der Gestalt der griechischen Helden alle Bewusstseinsfunktionen eingeschlossen waren, das Pferd selber noch längst kein Bewusstsein erhielt.

Bewusstseinsfunktionen, mechanisch oder nicht mechanisch, erzeugen kein Bewusstsein, solange zwei weitere entscheidende Dinge fehlen. Erstens die operative Einheit der Funktionen und zweitens die "transzendente" Steuerung dieser Einheit. Diese beiden zusätzlichen Funktionen sind leicht erklärbar.

⁹ Ein Existenzurteil ist ein Satz, der folgende Form hat: "Es gibt ..."

Übrigens schließt die Fähigkeit zum Existenzurteil nach der Formel

$\sim[(\text{Ex}) \sim f(x)] \equiv (x) f(x)$ die Fähigkeit zum All-Urteil ein.

"Alle ... " Hier unterscheidet sich der "mechanical brain" vom tierischen Bewusstsein. Es scheint, dass Tiere nur die Fähigkeit zum Existenzurteil haben, aber beim Vollzug von Allurteilen, wie Experimente zeigen, gewöhnlich versagen.

¹⁰ "Random element".

Wir besitzen heute Maschinen, die logische Operationen vollziehen, und andere Maschinen, die arithmetische Probleme lösen und algebraische Fragen und Aufgaben aus dem Gebiet der Infinitesimalrechnung beantworten. In beiden Fällen rechnen und arbeiten wir mit den verschiedenen Maschinen. Wir sind aber vorläufig nicht imstande, beide Maschinen so miteinander zu "koppeln", dass die Logikmaschine uns vertritt und an unserer Stelle mit dem arithmetischen Mechanismus arbeitet. In dem Moment, wo eine solche "Koppelung" glückte, wäre die erste Bedingung für ein mechanisches Bewusstsein, nämlich die operative Einheit aller seiner Funktionen gegeben. Was aber auch dann noch fehlte, wäre die "transzendente" Steuerung dieser gekoppelten Systeme.

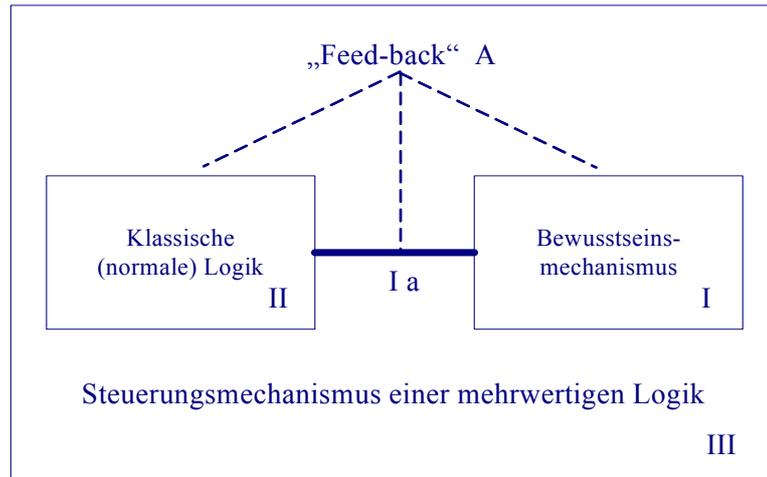
Dieser Steuerungsmechanismus müsste wieder eine Logik sein, aber eine, die einen höheren logischen Typus repräsentiert als unsere "normale" klassisch-aristotelische Logik, die mit der arithmetischen Maschine gekoppelt ist. Eine solche Logik eines höheren Typus ist keineswegs eine bloße Phantasie und ein metaphysischer Traum eines wildgewordenen Philosophen. Bruchstücke solcher Logiken, die unsere "normale" Logik als engen Spezialfall enthalten, existieren schon jetzt. Man nennt sie mehrwertige^[11] Logiken, und die theoretischen Mittel, sie weiter zu entwickeln, sind längst vorhanden.

Auch über die Arbeitsweise dieser "transzendenten" Steuerung unserer normalen Logik samt ihrer Inhalte lässt sich bereits einiges sagen.

Der transzendente Steuerungsmechanismus reflektiert Information auf sich selbst durch einen sogenannten Rückkoppelungskreis (feedback-system). Das klingt wieder schwieriger, als es in Wirklichkeit ist. Es heißt nichts weiter, als dass die folgenden mechanischen Schritte vollzogen werden. Erstens: der gekoppelte Mechanismus produziert Informationen. Zweitens: dieselben werden auf geeignete Weise an den Steuerungsmechanismus weitergeleitet. Da diese Steuerung aber selber eine Logik mit neuen Gesetzen ist, so verändert sie die an sie gelieferte Information (so wie sich für uns der Anblick der Welt verändert, wenn wir eine farbige Brille aufsetzen) und gibt dann drittens: ihren veränderten Gesichtspunkt als Steuerungsimpuls an das untergeordnete System zurück.

Damit sind alle Bedingungen für die Konstitution von Bewusstsein erfüllt. Wir besitzen dann ein System, das die materiellen Bewusstseinsmechanismen liefert. Ein zweites, das ihre operative Einheit herstellt, und ihnen übergeordnet ein drittes, das das gegenseitige Verhältnis des ersten und zweiten Systems reflektiert und das Resultat dieser Reflexion als Information in das Verhältnis von eins und zwei zurückkehren lässt. Diese Rückkehr der Information aber, die durch ein logisch überlegenes (und in das potentielle Bewusstsein nicht eingeschlossenes) System hindurchgegangen ist, ist derjenige Prozess, der *aktuelles* Bewusstsein erzeugt. Jedenfalls hat sich noch niemand unter Bewusstsein etwas anderes vorstellen können als Information (Erlebnissinn), die sich in einer sinnhaft modifizierten Weise auf sich selber bezieht und dadurch von sich selber weiß. Ein grobes Schema eines solchen "mechanical brain" hat also die folgende Gestalt:

¹¹ [Anmerkung_evgo]: Man spricht heute besser von einer "mehrstelligen" Logik



In diesem Schema enthält das Subsystem (I) alle Bewusstseinsmechanismen wie "Wahrnehmung", "Erinnerung", "Abzählen" usw.; also Mechanismen, die das aus der Außenwelt gelieferte Material der "Bewusstseinsinhalte" verarbeiten bzw. speichern (Erinnerung). Subsystem (II) integriert diese Funktionen im Sinn der klassisch-aristotelischen Identitätslogik, so dass sich jeder Bewusstseinsmechanismus auf ein und dieselbe Identität in seinen Informationsdaten bezieht. Dies soll an einem Beispiel erläutert werden: Ein beliebiges Bewusstsein, sei es tierisch, menschlich oder mechanisch, das alle überhaupt möglichen Informationsdaten in Form von Wahrnehmung, Erinnerung usw. über ein Krokodil besitzt, hat darum noch immer keinen Gesamtbegriff von einem solchen Tier. Um das zu erreichen, müssen erst alle Daten im Sinn einer generellen Identität "Krokodil" geordnet und in einer Einheit als objektiver Gegenstand integriert werden. Das ist die Aufgabe, die Subsystem (II) erfüllt. Nur teilweise allerdings, denn wir wissen aus der Beobachtung unseres eigenen Bewusstseins, dass hinter diesem Identitätssystem unseres eigenen Denkens ein rätselhaftes X steht, das wir Ich nennen. Dieses Ich folgt eigenen, anderen Gesetzen, die nichts oder nur wenig mit denen der klassischen Logik zu tun haben. Wir wissen aber, dass es die Beziehungen zwischen den Subsystemen (I) und (II) steuert. Diese Steuerung ist nicht eine einmalige Sache, sondern erfolgt auf vielen Operations- bzw. Handlungsebenen zugleich. Wir wollen die sukzessive Ordnung dieser Handlungsebenen "A", "B", "C", "D" usw. nennen.

Unser System (III) repräsentiert die erste oder unterste Operationsebene, auf der die Systeme (I) und (II) "transzendent" gesteuert werden. Diese Steuerung geschieht vermittelt eines Regelkreises (feed-back-system) von der Ordnung "A"^[12] Weitere Rückmeldungen (feed-backs) von der Ordnung "B", "C", "D" usw. sind in unserem System III nicht aktivierbar. Sie sind auch nicht nötig, insofern als Rückmeldung A bereits genügt, um alle Bedingungen zu erfüllen, die das elementare Phänomen von Bewusstsein (nicht Selbstbewusstsein) voraussetzt.

Wir haben keine Ursache, anzunehmen, dass, wenn alle mechanischen Bedingungen für das Zustandekommen von Bewusstseins-handlungen erfüllt sind, dann nicht

¹² Dass das menschliche Gehirn einen Regelkreis (feed-back-system) enthält, ist bereits festgestellt.

auch Bewusstsein (von Gegenständen) wirklich eintritt. Die Annahme, dass ein Ereignis nicht eintritt, wenn alle Bedingungen seines Eintretens erfüllt sind, widerspricht allen unseren wissenschaftlichen Erfahrungen. Wir nehmen also an, dass, sobald die Rückmeldungsoperation "A" in Bewegung gesetzt ist, Bewusstsein tatsächlich eintritt, und fragen, wo das geschieht. Die Antwort muss lauten – und sie gründet sich auf die heute bereits wohlbekanntes mechanischen Eigenschaften eines Rückmeldungsmechanismus – in einem Zwischensystem zwischen (I) und (II), das wir in unserer Zeichnung als Subsystem (Ia) bezeichnet haben.

Damit lässt sich einiges über die Eigenschaften eines solchen künstlich erzeugten Bewusstseins sagen. Sein Erlebnisraum wird ausschließlich Daten seiner ihm zugeordneten Außenwelt enthalten. Das heißt Daten, deren Ursprung außerhalb des Systems III liegt. Nicht einbezogen in das Erlebnispotential eines solchen Bewusstseins dagegen sind die Systeme (I) und (II). Das heißt, dieses Robotgehirn "weiß" weder, dass es Mechanismen hat, noch ist ihm erfahrbar, von welcher Logik seine Mechanismen integriert werden. Hier stoßen wir auf den entscheidenden Unterschied zum menschlichen Bewusstsein. Wenn es für uns nicht möglich wäre zu erfahren, dass in unserem Bewusstsein Mechanismen am Werke sind, dann könnten wir weder Psychiatrie noch Psychologie betreiben. Und wenn Aristoteles nicht ein Bewusstsein besessen hätte, dessen Erlebniskapazität das Subsystem (II) umfasste, dann hätte er niemals die nach ihm benannte Logik schreiben können. Ein Robot, der nach dem obigen Schema gebaut ist, wird niemals ein Buch über die Theorie seines eigenen Denkens schreiben können. Denken wird er zwar – in rein gegenständlichen Kategorien über die Außenwelt. Aber der obige Mechanismus enthält keine Möglichkeiten, die ihm erlauben würden, über sein eigenes Denken nachzudenken. Um einen psychologischen Ausdruck zu gebrauchen: ein solcher Robot wird ein hundertprozentiger Extrovert sein. Das heißt, er kann nicht über sich selber nachdenken, weil ihm die Begegnung mit dem Ich-Erlebnis fehlt. Der hier geschilderte Robot hat zwar Bewusstsein (von Gegenständen), aber kein Selbstbewusstsein. Das einzigartige Erlebnis des Ichs oder des Selbstes im Selbstbewusstsein bleibt für alle Ewigkeit dem Menschen vorbehalten und kann von ihm in keinem in eine Maschine projizierten Entwurf je dupliziert werden.

Kant sagt in seinem berühmten Satz in der Kritik der reinen Vernunft: "Das: Ich *denke* muss alle meine Vorstellungen begleiten *können* ... " Das Ich-Erlebnis (Selbstbewusstsein) braucht also keineswegs die Bewusstseinsakte faktisch zu begleiten, aber es muss potentiell hinter ihnen stehen. Wie verhält es sich damit im Fall eines "mechanical brain", dessen prinzipielle Beschränkungen wir eben aufgezeigt haben? Allgemein muss dazu bemerkt werden: Was hier für den Menschen gilt, gilt für ein Robotgehirn erst recht, wenn es Bewusstsein besitzen will. Nur ist das Ich im Fall des Robot aus dem Mechanismus in den Konstrukteur zurückverlegt. Er repräsentiert das "Ich" des Mechanismus. Ein sehr wesentliches Element der kybernetischen Theorie ist nämlich, dass die Konstruktionsideen des Ingenieurs, der den "mechanical brain" entwirft, mit dem Robotgehirn zusammen zwar kein physisches, wohl aber ein logisches System bilden. Mit anderen Worten: Um erfolgreich einen "mechanical brain" zu entwerfen, muss sich das Bewusstsein des Konstrukteurs in einem Regelkreis (feed-back-system) von der Ordnung "B" an das Robot-System "A" direkt anschließen. Und diese Systeme "A", "B", "C", ...

setzen sich in Richtung auf die Transzendenz des Ichs in einer unendlichen Serie fort. Nicht nur dem Robot ist sein "Ich" transzendent und unerreichbar; das gleiche gilt auch für die Seele des Konstrukteurs.

Hier ist der naturalistische Einwand zu erwarten: Wie kann ein Robot durch die Vermittlung des Ichs des Konstrukteurs denken! Denn das letztere ist nicht sein eigenes, sondern ein fremdes Ich. Darauf kann nur geantwortet werden: Wenn nicht die Metaphysik aller Völker und Zeiten gelogen hat, dann sind alle Iche, ontologisch betrachtet, identisch. Niemand kann angeben, was, wenn man das Ich nach einer unendlichen Regression endlich erreicht, der Unterschied zwischen "eigen" und "fremd" eigentlich sein soll. Das ist eine Unterscheidung, die sich passender auf Zahnbürsten anwenden lässt. Wem aber metaphysische Argumente keinen Eindruck machen – weil er noch nicht gelernt hat, Metaphysik von Mythologie zu unterscheiden – braucht ja an Iche, Seelen oder Subjekte und ihre metaphysische Identität überhaupt nicht zu glauben. Es ist aber eine ganz empirische Erfahrung, dass alle Subjektivität "bodenlos" ist. Das heißt, es liegt hinter jedem erreichten Bewusstseinszustand immer noch ein tieferer, nicht erreichter. Und jedes überhaupt mögliche Bewusstsein bescheidet sich dabei, auf diesem Weg ins Innere an einer Stelle haltzumachen. Der Halt ist willkürlich, aber das liegt im Wesen der Sache und kann nicht geändert werden.

Diese "Bodenlosigkeit" aber ist in dem von uns skizzierten Robotgehirn, *relativ zu der dort erreichbaren Bewusstseinsstufe*, bereits etabliert. Sowohl die Subsysteme (I) und (II) wie der Rückmeldungsmechanismus liegen erlebnismäßig "hinter" dem Robotbewusstsein. *Wir* wissen, dass das Robotbewusstsein nicht bodenlos ist und schnell seinen Boden in den Grenzen des Systems III erreicht. Aber der Robot kann das nie wissen. Die spezifizierte Konstruktion seines Bewusstseins schließt eine solche Erkenntnis für immer aus. Und es darf nicht vergessen werden, dass der Mensch hier in der gleichen Lage ist. Wir sprechen zwar von dem infiniten Regress der Subjektivität in sich selbst. Das kann aber, wenn wir uns um Präzision in dem Ausdrucksgehalt unserer Aussagen bemühen, nichts anderes heißen, als dass unser Bewusstsein so organisiert ist, dass wir nie seinen Boden erreichen können. Einen Boden, der vielleicht schon unmittelbar jenseits der gerade erreichten Reflexionsschwelle liegt. Machen wir eine weitere Reflexanstrengung, so weicht der Boden genau um einen Schritt zurück. Unser Bewusstseinsystem wäre in diesem Falle endlich, und doch könnten wir sein Innerstes, d.h. seinen Boden, nie erreichen. Überdies besitzt das Robotbewusstsein eine echte Verbindung mit der Realität, die über die Grenzen seines Systems III hinausgeht. Nur tritt dieselbe nicht in Erscheinung, wenn man sie vom entworfenen Mechanismus her aufzuspüren versucht. Sie wird nur vom Konstrukteur der Maschine her sichtbar. Und damit sind wir an den Ausgangspunkt unserer Betrachtung zurückgekehrt.

Wir gingen von der metaphysischen Bedeutung der menschlichen Technik aus und bemerkten auf den einleitenden Seiten dieses Kommentars, dass der Mensch in der Maschine sich einen zweiten "Leib" schafft, in den er seinen physischen Arbeitsschematismus hinein projiziert hat. Warum aber geschieht das? Die Antwort darauf ist vage bei Denkern wie Hobbes, Vico und Fichte antizipiert, in scharfer Formulierung mit dem Bewusstsein aller Konsequenzen aber erst in der amerikanischen

Philosophie des Pragmatismus gegeben worden. Der Mensch versteht nur das, was er macht. Alle Bewegung eines freibeweglichen Organismus ist essentiell ein Verstehensprozess. Aus diesem Grunde besitzt schon das Tier Verstehens-Kategorien, da es sich willkürlich bewegen, also handeln kann. Aber die Handlungen, die ein freibewegliches Wesen ausschließlich mit seinem Körper ausführt, liefern nur Verstehenskategorien für die Außenwelt, die diesen umgibt. Der Körper selbst und das in ihm wohnende "Leben" bleibt in einer einfachen Bein- und Rumpfbewegung unverstanden. Diese Bewegungen sind "instinktiv", d.h. sie reflektieren nicht auf sich selber. E. T. A. Hoffmanns "Kater Murr", der über sich selbst reflektiert, existiert in der Tierwelt nicht.

Wenn der Mensch sich selbst – und um damit irgendwo einen Anfang zu machen – seinen Körper als freibewegliche Existenz innerhalb der Umwelt verstehen will, bleibt ihm nichts anderes übrig, als diesen Körper als Maschine zu wiederholen. Wir wissen zwar auch ohne Technik, *dass* wir laufen, aber *wie* wir laufen, verstehen wir erst dann adäquat, wenn uns der Mechanismus von Gelenk und Hebel kein Geheimnis mehr vorenthält.^[13] Nun dehnt sich aber der Handlungsbereich jedes freibeweglichen Wesens weit über den Aktionsradius rein körperlicher Handlungen aus. Fichte hat zuerst – auf deduktivem Wege – überzeugend nachgewiesen, dass Bewusstsein als eine Tätigkeit des Menschen verstanden werden müsse, wenn man es überhaupt verstehen will. Damit ergibt sich als selbstverständliche Forderung: Will der Mensch sein eigenes Bewusstsein und dessen Prozess verstehen, so bleibt ihm nicht anderes übrig, als dasselbe als Handlung, d.h. in einem technischen Herstellungsverfahren in der Außenwelt zu wiederholen. Mit Introspektion ist hier nichts zu machen.

Nun weiß jeder, der etwas vom Arbeitsrhythmus des Körpers und von Maschinenteknik versteht, dass nicht die ganze Bewegungsschematik des lebendigen Tieres oder Menschen auf die Maschine übertragen wird. Es ist immer nur ein äußerst geringer Bruchteil. Am Prinzipiellen der Sache aber ändert sich dadurch nichts: es werden Ereignisfolgen, wie sie die "Natur" aus sich niemals allein entwerfen könnte, aus der handelnden Subjektivität partiell abgelöst und auf den objektiven Seinszusammenhang transferiert, um dort eine autonome Existenz zu führen. Das eilt für alle vergangene Technik. Diese Interpretation aber trifft ebenso auf eine zukünftige nicht-klassische Robottechnik zu. Der metaphysische Hintergrund des technischen Bemühens bleibt derselbe, und die Konstruktion eines "mechanical brain" ist nicht rätselhafter oder unmöglicher als die Konstruktion eines Karrens mit Rädern. Letztere würde dem adamitischen und noachitischen Menschen, wenn man versucht hätte, ihm zu erklären, dass man das Voreinandersetzen der Füße mechanisch in eine rotierende Bewegung von unendlicher Kontinuität umsetzen könne, als ein Wunder erschienen sein, dessen Erfüllung sich notwendig die Götter vorbehalten hätten.

Eine wissenschaftliche Formulierung der Aufgabe, wie sie in den kybernetischen Theorien erfolgt, beabsichtigt nun nichts anderes, als was der Mensch getan hat, als er seinen ersten Wagen baute, nämlich die Abspaltung eines partiellen Aus-

¹³ Es ist kein Zufall, dass das Radschlagen im Turnunterricht eine halbreflektierte Körperübung ist. Tiere spielen zwar, aber sie turnen nicht.

schnittes aus dem Handlungsbereich des Bewusstseins und Übertragung dieses Ausschnittes auf isoliert objektive Seinszusammenhänge. Es handelt sich also gar nicht darum, eine "Seele" zu konstruieren. Wer den letzteren Vorschlag wirklich machen wollte, gehörte unter ärztliche Beobachtung. Wohl aber ist es eine ernsthafte wissenschaftliche Frage, ob es möglich ist, partielle Bewusstseinsräume vom tierischen bzw. menschlichen Organismus abzuspalten und auf einen (trans-klassischen) Mechanismus zu übertragen.

Die Frage wird heute in Amerika zögernd bejaht (Edmund C. Berkeley), da unter pragmatistischen Gesichtspunkten an einer solchen technischen Prozedur nichts wunderbarer ist, als wenn man einen sinnvollen Arbeitsrhythmus (wie ihn die Natur nie allein hervorbrächte) vom menschlichen Körper ablöst und auf Räder und Hebel überträgt. In gleicher Weise spaltet der Konstrukteur eines "mechanical brain" einen partiellen Funktionsbereich seines Bewusstseins ab und überträgt dessen Arbeitsrhythmus auf elektrische Relais und nucleare Vorgänge in der Außenwelt. Unser klassisches Gefühl sträubt sich heute noch gegen diesen Parallelismus, weil wir uns zu dem Einwand getrieben fühlen: ein "mechanical brain" muss bewusstlos bleiben, da der Konstrukteur doch nichts von *seinem* Seelenleben abspalten kann.

Dem muss erwidert werden, dass der Dualismus von Leib und Seele in der hier skizzierten Betrachtungsweise völlig gegenstandslos ist. An dieser Stelle kommt die deutsche philosophische Anthropologie den kybernetischen Theorien Amerikas zu Hilfe. In seinem tiefgründigen Werk "Der Mensch" (Bonn 1950) hat Arnold Gehlen ganz überzeugend demonstriert, dass die menschliche Wirklichkeit – unter Einschluss des Bewusstseins! – in einem Begriffssystem beschrieben werden kann, das den Unterschied von Leib und Seele nicht mehr kennt. Voraussetzung dafür allerdings ist, dass man von der Grundvoraussetzung ausgeht, dass der Mensch seine Wirklichkeit nur als Handelnder erfährt und versteht. Die Gesetzlichkeit eines solchen universalen Handlungsvollzuges, der unser Bewusstsein einschließt, ist in der Tat das Grundthema der kybernetischen Forschung.^[14]

Das Problem, wie Bewusstsein entsteht, erhält in dieser kybernetischen Denkweise eine neue Formulierung. Gehlen sagt dazu: Bewusstsein bedeutet einen Umsetzungsvorgang (uns übrigens undurchsichtiger Art) an den Kontaktstellen eines Organismus mit der Welt.^[15] Diese Auffassung kann als technisches Problem sehr präzise formuliert werden: Es gilt ein geschlossenes System (Äquivalent zu Organismus) zu konstruieren, das regulierte Kontaktstellen mit der Außenwelt besitzt. Jeder solche Kontakt muss als Information verarbeitet werden können. Und das geschlossene System muss ein Informationssystem von in sich reflektiertem Charakter sein. — Die Reflexion der Information auf sich selbst (der eigentliche Bewusstseinsvorgang) kann dadurch hergestellt werden, dass man die Information durch zwei verschiedene logische Systeme gehen lässt, wobei das zweite System die Bedeutung, die das Informationsmaterial im ersten logischen System gehabt hat, in bestimmter Weise verändert und diesen veränderten Sinn an die erste Logik

¹⁴ Vgl. dazu Helmut Schelsky, *Zum Begriff der tierischen Subjektivität*, Studium Generale (1950) III, 2/3, S. 102-116.

¹⁵ A. a. O. S. 279.

zurückgibt. Dieser geschlossene Reflektionskreis tritt dann aufs neue mit den Informationsdaten in Kontakt. Und dies ist der Punkt, wo – der Theorie gemäß – Bewusstsein entstehen sollte. Unter den Kontakten, die von der rückkehrenden Reflexion erreicht werden, ist nämlich ein vor allen andern ausgezeichneter. Wir wollen ihn den Initialkontakt nennen. Zum Informationsbestand der Maschine gehören nämlich einige "metaphysische" Daten: erstens die Information, *dass* der Konstrukteur den "mechanical brain" geschaffen hat und *wie* er ihn geschaffen hat, und weiter: *dass* er ihn in Bewegung gesetzt hat und *wie* er ihn in Bewegung gesetzt hat. Ohne die Verarbeitung dieser Information wäre der "mechanical brain" nie in der Lage, einen Analogprozess zu menschlichem Bewusstsein zu produzieren.

Diese heute noch so unmöglich anmutende Produktion ist ja im Grunde nichts anderes als ein einfacher Transmutationsprozess "seelischer" Energien. In der Schöpfung und Ingangsetzung eines "mechanical brain« setzt der Konstrukteur *sein* Bewusstsein in Handlung um. Eine Handlung aber ist kein Bewusstsein mehr, obwohl sie von Bewusstseinsakten begleitet sein kann. Ein "mechanical brain" nun ist seiner Idee nach ein äußerst verwickelter Mechanismus, der diesen Prozess der Umsetzung bewusster Reflexion in nicht bewusste Handlung partiell wieder rückgängig macht. Wir sagen ausdrücklich "partiell", denn zur Handlung gehört auch das Handlungsergebnis, d.h. das fertiggestellte Robotgehirn, das ja durch diesen Prozess nicht berührt werden darf, denn sonst würde es sich in Rauch, Ideen und Rohmaterial auflösen. Abgesehen vom Handlungsergebnis aber ist auch die Handlung selbst in der Maschine (als Information) enthalten, und es ist nicht einzusehen, warum der Mensch, wenn er einmal sein Bewusstsein in Handlungen transformiert hat, nicht vermittels eines Hilfsmechanismus ("mechanical brain") diesen Prozess dann reversieren und Handlungen in Bewusstsein zurückverwandeln kann.

Es geschieht also gar nichts Mystisches im Robotgehirn, und dasselbe hat auch gar kein "eigenes" Bewusstsein. Wenn die hier geschilderten Ideen sich wirklich durchführen lassen, so würde das nichts anderes bedeuten, als dass es dem Menschen gelungen ist, Bewusstseinsvorgänge partiellen Charakters von seinem Organismus abzulösen und auf ein anderes Medium zu übertragen. Ein Mechanismus *erzeugt* kein Bewusstsein, auch nicht, wenn sein Arbeitsrhythmus trans-klassisch ist. Eine ganz andere und völlig offene Frage aber ist, ob es ihm nicht möglich wäre, Bewusstseinsakte, die man in ihm in Handlungsformen deponiert hat, wieder in ihren ursprünglichen Zustand zurückzuführen. Bewusstsein ist keine stoffliche Quantität, die erzeugt oder vernichtet werden kann, sondern eine metaphysische Variante von Existenz, die, wie das Verhältnis von Denken und Wollen zeigt, sich willig aus einem existenziellen "Aggregatzustand" in einen anderen überführen lässt. In dem hier vorliegenden Fall würde der Mensch durch seine *eigene*, in der Maschine nur investierte Handlung einen Teil seiner Handlungskapazität in Intelligenz zurückverwandeln. —

Die Robots, die Isaac Asimov in seinen Erzählungen einführt, entsprechen dem hier geschilderten Bild eines "mechanical brain" nicht ganz. Der Autor hat ihnen mit dichterischer Freiheit mehr Intelligenz gegeben, als das nach heute maßgeblichen logischen Theorien, auch unter Voraussetzung einer neuen trans-klassischen

Naturwissenschaft, möglich wäre. Um die Intelligenzleistungen Asimovscher Robots zu erzielen, würden wir "mechanical brains" mit einem dreifachen Regelkreis (feed-back-system) benötigen. Der erste Rückmeldungstransport würde, wie oben beschrieben, durch eine dreiwertige Logik führen. Der nächste durch eine vierwertige und der dritte durch ein fünfwertiges System des Denkens. Es gibt aber und wird auch für lange zukünftige historische Entwicklungen kein menschliches Gehirn geben, das sich auch nur im geringsten vorstellen kann, *was* man mit vier- oder gar fünfwertigen Logiken denken kann. Man kann zwar solche hypothetische Systeme mechanisch auf dem Papier konstruieren, aber niemand weiß, was die dann entstehenden Formeln inhaltlich eigentlich meinen.

Eine weitere Differenz zwischen wissenschaftlicher Theorie und poetischer Lizenz kann der Leser in der dritten Geschichte "Vernunft" finden. Da in das Robotgehirn Information über seinen Schöpfer eingeschlossen sein muss, da andernfalls, Denkprozesse nicht entstehen können (es ist ja nur von dem Konstrukteur geborgtes Denken), ist es technisch ganz unmöglich, dass der Robot QT-1 die Gedankengänge entwickelt, die er dort vorträgt. Sie würden der aus dem "Initialkontakt" abgeleiteten Information widersprechen. Ihre Produzierung wäre also mechanisch unmöglich.

Das alles aber ist unwesentlich, denn das gesamte Asimovsche Buch ist einem speziellen Problem gewidmet, das in der Tat schon bei dem Typ von "mechanical brain" auftaucht, den wir als wissenschaftlich möglich beschrieben haben. Asimovs Thema ist: eine Maschine, die Bewusstsein hat, braucht einen Moralkodex. Mehr noch: sie kann technisch ohne einen solchen überhaupt nicht konstruiert werden. Denken (auch geborgtes) setzt Spontaneität voraus, und diese braucht *ethische* Richtlinien. Wenn man aber Ethik in eine Maschine einbaut, dann entstehen allerhand Schwierigkeiten – zu denen die einzelnen Geschichten sehr amüsante Illustrationen liefern.

Man lasse sich durch den leichten Ton nicht täuschen. Es steckt viel Tiefsinn in den elegant vorgetragenen Episoden. Und der Verfasser, der Professor der Biochemie an der Universität Boston in Massachusetts (USA) ist, weiß sehr genau, dass seine Robots dem, was in den nächsten Jahrhunderten wissenschaftlich möglich sein wird, noch um Erhebliches voraus sind. Um das anzudeuten, hat er seine Robots mit einer speziellen technischen Zukunftsphantasie, dem "positronic brain" ausgestattet. Niemand weiß heute, wie ein solcher Asimovscher "positronic brain" eigentlich funktionieren sollte. Der Autor weiß es selbst nicht. Die Idee dieses neuen "mechanical brain" hat nur den Rang eines Symbols, das andeuten soll, dass eine neue Kultur erst eine neue Idee der Wissenschaft und der ihr folgenden Technik entwickeln muss, ehe Schöpfungen wie denkende Robots möglich sein werden. Von hier aus gesehen sind die Asimovschen Ideen kaum zu kritisieren. Nur die letzte der Geschichten Der unvermeidbare "Konflikt" muss der Wissenschaftler genau so wie der Philosoph beanstanden. In dieser Erzählung dankt der Mensch historisch ab und die Leitung der menschlichen Geschichte wird von den Robots übernommen. Der im historischen Denken geschulte Europäer wird über diese utopische Perspektive lächeln. Die letzten Konflikte menschlicher Geschichte liegen tiefer, als dass sie je auf der Existenzebene der Maschinen auftreten könnten. Die in der neunten und abschließenden Erzählung entwickelten Auffassungen über Ge-

schichte sind aber symptomatisch für die generelle amerikanische Haltung gegenüber historischen Problemen. Der Amerikaner fühlt dunkel, dass alle bisherige Weltgeschichte im strengen Sinn eine Angelegenheit der östlichen Hemisphäre war und ihn im tiefsten Grunde nichts angeht. Alles aber, woran man nicht, oder nicht mehr, seelisch beteiligt ist, wird zu einem rein technischen Problem des Lebens. Eigene und neue historische Erlebniskategorien besitzt der Mensch der westlichen Hemisphäre noch nicht, weshalb seine Ideen über Geschichte den Gedankengängen eines Fabrikleiters ähneln, der den Produktionsprozess in seinen Werken verbessern will.

An dieser Stelle stoßen wir auf die letzten weltanschaulichen Motive, die allen amerikanischen kybernetischen Theorien ihre innere Triebkraft geben. Der neue Mensch der westlichen Hemisphäre identifiziert sich nicht mehr mit den reinen Formen des klassischen Denkens, die in einer langen und mühevollen Seelengeschichte des Menschen in der östlichen Hemisphäre entwickelt worden sind. Er sucht diese Formen dadurch von sich abzustößen und sie innerlich zu Überwinden, dass er versucht, sie aus seinem Seelenleben zu entlassen und in die Maschine, den denkenden Robot, zu verbannen.

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

Copyright 2004 vordenker.de

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

vordenker

ISSN 1619-9324