Sitzungsberichte der Leibniz-Sozietät der Wissenschaften zu Berlin *113(2012), 23–46*

Helmut Moritz

Über G. Chaitin – Von Metamathematik zur Metabiologie (Ein Beitrag zur Wirkung Leibnizscher Ideen)

Vortrag im Plenum der Leibniz-Sozietät am 13. Oktober 2011

Dem Gedenken an Hans-Jürgen Treder und Wilfried Schröder gewidmet

Zusammenfassung

Im Gegensatz zur *Physik*, die seit Newton wesentlich und unzertrennlich mit der Mathematik verbunden ist (eben durch die Differentialrechnung von Leibniz und Newton), ist die *Biologie* bisher nicht so selbstverständlich und durchgehend mathematisiert. Gregory Chaitin versucht, seit 2009, eine wesenhafte und durchgehende axiomatische Mathematisierung der biologischen Evolution. Dies geschieht allerdings nicht durch die herkömmliche Mathematik, sondern durch die Metamathematik, wie sie von Russell, Hilbert, Gödel und Turing entwickelt wurde. Letzterer hat einen wesentlichen Zusammenhang mit der Computertheorie aufgezeigt.

Die Biologie hat einen kontinuierlichen Aspekt und eine diskrete Struktur. Letztere ist der genetische Code, der für die Evolution grundlegend ist. Wenn wir also hier von Biologie sprechen, meinen wir die *Evolution des genetischen Codes*.

Stichworte für diese Entwicklung sind Hilberts Entscheidungsproblem, Gödels Unvollständigkeitstheorem und der idealisierte Turing-Computer. Zusammen mit der Kolmogorow-Komplexität und der von Chaitin u.a. entwickelten Theorie der algorithmischen Informationstheorie ergibt das die Chaitinsche Auffassung, dass die Theorie der binären Folgen mit dem „Alphabet“ (1, 0), wie sie für die Computer-Theorie grundlegend sind, ein sehr vereinfachtes aber umfassendes und tief liegendes „Spielzeugmodell“ (toy model) für die biologische DNA mit dem bekannten Alphabet (A, G, C, T) darstellt. Nach Chaitin ist DNA also die „universelle Computersprache der Biologie“. Den unbekannten exakten Theoremen der DNA-Sprache entspre-

chen einfachere, aber nicht minder exakte, Theoreme einer binären „Computer-Sprache“, die Chaitin betrachtet.

So kann Evolution angesehen werden als Zufallsbewegung (*random walk*) in einem Raum von binären Folgen. Die biologische Fitness entspricht der Fähigkeit, sehr schnell wachsende Folgen von großen Zahlen zu erzeugen.

Die Biologie ist gewiss hoch komplex, aber diese Komplexität hat ein mathematisches Gegenstück in der schon relativ gut bekannten binären algorithmischen Informationstheorie. Nach Waldrop und Ebeling liegt Komplexität „an der Grenze zwischen Ordnung und Chaos“. In der neuen Theorie kommt das durch die inzwischen wohlbekannte Chaitin-Konstante zum Ausdruck, welche diesen beiden Aspekten Ordnung und Chaos, oder nach Monod, Zufall und Notwendigkeit, entspricht.

Chaitins Arbeit ist noch keineswegs abgeschlossen, stellt aber ein viel versprechendes „work in progress“ dar, das seit 2009 bereits wesentliche Ergebnisse gezeigt hat.

Abstract

Contrary to *physics*, which since Newton has been essentially and inseparably connected with mathematics (by the differential calculus of Leibniz and Newton), the relation of biology to mathematics has been by far not so thoroughgoing and self-evident. Gregory Chaitin, since 2009, tries to establish an essential and natural connection between biological evolution and mathematics. However, this mathematization of biology is not accomplished by classical calculus, but by metamathematics as established by Russell, Hilbert, Goedel and Turing. It was Turing who recognized an essential connection with the theory of computation.

Biology contains a continuous aspect and a discrete structure. The latter is the Genetic Code, which will be exclusively considered in this paper. Whenever we speak of biology, we mean *evolution of the genetic code.*

Keywords for this development are Hilbert’s decision problem, Goedel’s incompleteness theorem, and the halting problem for an idealized Turing computer. Together with Kolmogorov’s definition of complexity and Chaitin’s theory of algorithmic information, this results in Chaitin ingenious idea that the theory of binary strings with basis or “alphabet” (1, 0), fundamental for theoretical computation, gives a very simplified but comprehensive and profound “toy model” for the biological DNA with the well-known alphabet (A, G, C, T). According to Chaitin, DNA thus provides a “universal computer

language of biology”. To the unknown exact theorems of the DNA language, there correspond simpler, but not less exact, theorems of a binary “computer language” as considered by Chaitin.

Thus, evolution may be regarded as a random walk in a space of binary sequences. Biological fitness is the capability of producing very rapidly increasing sequences of large numbers.

Biology certainly is highly complex, but this complexity has a rather well understood mathematical counterpart in the binary theory of algorithmic complexity. According to Waldrop and Ebeling, complexity is situated “at the edge between order and chaos”. In the new theory this is particularly well expressed by the now well established Chaitin constant, which corresponds to these aspects of order and chaos, or, according to Monod, of randomness and necessity.

Chaitin’s work is by no means finished, but it represents a very promising “work in progress”, which so far has already provided essential results.

1. Einführendes über Chaitin und mathematische Biologie

Schon seit meiner Hochschulzeit interessierte ich mich für Philosophie und Grundlagen der Mathematik. Als Professor hielt ich Vorlesungen über Philosophie der Naturwissenschaften, die ich im Buch „Science, Mind and the Universe“ 1995 [20] veröffentlichte. Nachdem das Buch vergriffen war, stellte ich es in meine Webseite www.helmut-moritz.at, von wo es kostenlos heruntergeladen werden kann.

Dass das Buch Anklang fand und noch immer als aktuell gilt, zeigt eine 2010 erschienene russische Übersetzung durch V. Abalakin, den ehemaligen Direktor der berühmten Sternwarte in Pulkowo.

Nachdem ich mich, auf Anregung meiner Tochter, etwas mit Biologie beschäftigt hatte [18], auch mit dem Zusammenhang von Biologie und Mathematik, doch unzufrieden mit Literatur über „mathematische Biologie“, fand ich im Internet 2009 eine Arbeit von Gregory Chaitin [9], die mich sofort begeisterte und deren Fragestellungen ich in die englische Fassung meines Leibniz-Vortrags 2008 [21] aufnahm. Seit 2010 bin ich mit Chaitin in hochinteressanter E-Mail-Korrespondenz. Seine um 2000 bei Springer erschienenen sehr gut lesbaren Bücher über Philosophie der Mathematik, die ich seinerzeit als anregende Urlaubslektüre verwendet hatte, hatten mir ihn als genialen „Querdenker“, als Metamathematiker, gezeigt, z.B. [5]. Er stellte mit Begeisterung, Phantasie und literarischer Qualität höchst originelle, zum mathematischen Mainstream komplementäre Gedanken vor.

In Chaitin [9] habe ich eine brillante Aufzählung der Probleme gefunden, die ich jetzt deutsch zitiere.

Aufgaben und Ziele nach Chaitin

- *Wir wollen die fundamentalen mathematischen Ideen aufzeigen, die in der Biologie enthalten sind.*
- *Wir wollen Theoreme über extrem einfache unrealistische Modelle beweisen, nicht aber Simulationen von extrem komplizierten realistischen Modellen durchführen.*
- *Unser Ziel ist nicht eine realistische Simulation der biologischen Evolution, sondern eine mathematische Darstellung der fundamentalen biologischen Prinzipien, die der Evolution zu Grunde liegen, aus denen wir beweisen können, dass die Evolution stattfinden muss.*
- *Das kann man als ein Spielzeugmodell ("toy model") betrachten, aber wir sehen es nicht als Spielzeug, wir sehen es als einen Weg, von unwesentlichen Ablenkungen zu abstrahieren, die nur das Bild verkomplizieren.*
- *Theorien sind Lügen, die uns helfen, die Wahrheit zu erkennen (Picasso).*
- *Wenn Darwins Theorie der Evolution so fundamental und allgemein ist, wie die meisten Biologen denken, dann muss es möglich sein, einige grundlegende mathematische Ideen daraus zu gewinnen (wie das bei der Physik seit Newton der Fall ist).*
- *Nichts ergibt Sinn in der Biologie außer im Lichte der Evolution (Dobzhansky).*
- *Es ist ein Skandal, dass wir keinen mathematischen Beweis dafür haben, dass die Evolution funktioniert!*
- *Ich (G.Ch.) bin ein reiner Mathematiker, kein Biologe: ich versuche, die platonische Idee der Evolution zu finden, das archetypische Verhalten, nicht die komplizierte (messy) Version, die in der wirklichen Welt stattfindet.*
- *Das Ziel sind Beweise, nicht realistische Simulationen.*
- *Eine weitere Art und Weise, in der sich unser Modell von früheren Modellen unterscheidet: Unser Ziel ist, die biologische Kreativität und die größeren Sprünge in der Evolution zu verstehen, nicht die allmählichen Veränderungen.*

  (Ende Zitat.)

Zur allgemeinen Geschichte der theoretischen Biologie siehe auch [25].

2. Komplexität, Korrelation und Prädiktion

Jedem von uns ist klar, dass ein Lebewesen etwas höchst Komplexes ist. Eine Ameise ist wesentlich komplexer als jeder Computer. Ein paar Beispiele wachsender Komplexität:

- Ein Sandhaufen kann zwar groß sein, besteht aber meist aus kleinen, lok-ker
- angehäuften, wenig zusammenhängenden Sandkörnern, den man wohl kaum
- als sehr komplex bezeichnen wird.
- Ein Ameisenhaufen ist schon deutlich komplexer, weil es sich um ein Bauwerk von Ameisen handelt (nicht zerstören!).
- Biologische Lebewesen wie Ameisen, Blumen oder Menschen sind natürlich hoch komplex strukturiert. Auch die Darwinsche Evolutionstheorie zeigt ein sehr komplexes Zusammenwirken von „Zufall und Notwendigkeit“ (Monod).

Ein solches Zusammenwirken kann am statistischen Begriff der *Korrelation* veranschaulicht werden, und zwar am geodätischen Beispiel des Erdschwerefeldes. (mein Fachgebiet). Dieses Feld wird beschrieben durch eine mathematisch exakte Theorie, die Potentialtheorie, hat aber auch eine statistische, oder eher komplexe, Struktur wegen der unregelmäßig-regelmäßigen Topographie der Erdoberfläche (wie alle Landschaftsmaler, etwa Caspar David Friedrich, wissen), wegen der Dichte-Unregelmäßigkeiten des Erdgesteins usw. Daher sind Schweremessungen, die etwa 1 km auseinander liegen, nicht gleich, werden aber auch nicht sehr voneinander verschieden sein. Eine solche Verschiedenheit kann nicht mathematisch-analytisch beschrieben werden, sondern nur statistisch, durch den wohlbekannten Begriff der Korrelation.

Eine *Vorhersage* oder *Prädiktion* verschiedener Werte des Erdschwerefeldes hat aber den Charakter einer rein analytischen Operation mit eindeutig definiertem Ergebnis, sobald man eine empirisch zu bestimmende Kovarianz-Funktion konsequent verwendet, welche die Korrelation der Größen des Erdschwerefeldes beschreibt; siehe [17], [19]. Die Bestimmung des Erdschwerefeldes auf diese Weise heißt *Kollokation;* sie beruht auf einem Hilbert-Raum.

Noch viel bekannter kommt dies in den *Wettervorhersagen* zum Ausdruck, die auf den Differentialgleichungen der Meteorologie beruhen, die aber auch als wesentliches Element die heute sehr bekannte *Chaostheorie* enthalten [21].

Waldrop [29], Ebeling et al. [13] und andere sagen, dass Komplexität an der Grenze von (*at the edge between)* Ordnung und Chaos liegen [20], S.171].

Noch ausgeprägter ist das statistische (*stochastische*) Element in der *Erdbeben-Vorhersage*, wie sie die Gruppe um V. I. Keilis-Borok mit beachtlichem Erfolg betreibt; siehe Internet.

Das unregelmäßige und doch wieder sehr regelmäßige Bild eines Baumes wird von der *deterministischen Chaostheorie* (*sic*!) im Sinne von Schuster [26] gut beschrieben. In diesen Bildern, wie *fast in der ganzen, vom Menschen unberührten Natur, gibt es keine geraden Linien,* mit Ausnahme der Lichtstrahlen und vielleicht der Kristallkanten [21]!

Schließlich noch ein einfaches Beispiel aus der Biologie des Menschen: eine *medizinische Diagnose* hat einen in unserem Sinne ganz ähnlichen prinzipiellen Charakter einer komplexen Prädiktion: im menschlichen Körper hängt Alles mit Allem zusammen, aber eben nicht rein deterministisch, sondern komplex: als Synthese von relativ „starren", „deterministischen" Elementen und „weichen" Elementen mit statistischen Unschärfen. Ein grundlegendes Element der komplexen Struktur einer jeden Zelle des Menschen (wie jedes Lebewesens) ist die gemeinsame DNA (Abschn. 5.1).

3. Chaitins algorithmische Informationstheorie

Einen Versuch, den hier informell beschriebenen Begriff der Komplexität u.a. für den Zweck einer Anwendung auf die biologische Evolution zu präzisieren, stellt die algorithmische Informationstheorie dar, die von Solomonoff 1960, Kolmogorow 1965 und Chaitin 1965-1975 begründet wurde; siehe den Wikipedia-Artikel (**WP**) "Kolmogorov Complexity".

3.1 Die algorithmische Komplexität

Wir folgen zunächst fast wörtlich der klaren Darstellung von Ebeling et al. [13], S. 25-26.

Betrachten wir eine binäre Folge oder „String" p der Länge $\ell(p)$:

$$p = 0000101010000101010111 \ (N = 22). \tag{1}$$

Kolmogorow verknüpfte seinen Komplexitätsbegriff mit einem binären Programm q, das durch einen Algorithmus oder einen Computer F die Folge p generiert, so dass $p = F(q)$. Eine Folge heißt zufällig oder maximal komplex, wenn es kein Programm q gibt, dessen Länge N kürzer ist als die der

Originalfolge *p*. Kolmogorow definierte als *algorithmische Komplexität* der Folge *p* die minimale Länge eines erzeugenden Programms

$$K_F(p) = \min N(q);\ F(q) = p. \tag{2}$$

Nach dieser Auffassung ist die Komplexität der periodischen Folge

$$p = 10001000100010001\ldots \tag{3}$$

gering, weil man für diese Folge ein kurzes Programm schreiben kann. Andererseits ist die Komplexität einer Folge, die durch das Werfen einer Münze entsteht, wie z.B.

$$p = 01010011010011011100\ldots \tag{4}$$

sehr groß, weil man kein Programm schreiben kann, das wesentlich kürzer als die Folge ist.

Aus mathematischer Sicht ist der Begriff der mathematischen Komplexität eine große Errungenschaft; aus praktischer Sicht wirft diese Definition jedoch eine Reihe von Problemen auf, unter anderem:

1. Es gibt keinen „Algorithmus“, mit dessen Hilfe man die algorithmische Komplexität allgemein berechnen könnte.
2. Es widerspricht unserer Intuition, dass eine zufällig generierte Folge, die etwa durch Münzenwerfen entstanden ist, maximal komplex sein soll.

Soweit das Zitat nach Ebeling. Diese beiden Argumente sind m.E. sehr tiefliegend, und wir werden später noch darauf zurückkommen (Kapitel 4). Im Übrigen formuliert Chaitin die Aussagen von Ebeling noch schärfer: „We can almost never determine the complexity of anything, if we chose to measure complexity in terms of the size of the smallest program for calculating it!“ [10], S .81). Und doch funktioniert der Begriff für unseren Zweck, wie wir in Kapitel 5 sehen werden.

Zusammenhang mit Wahrscheinlichkeit und Entropie

Wir können den Kolmogorowschen Komplexitätsbegriff leicht aus dem aus der Physik wohlbekannten Begriff der *Entropie* intuitiv ableiten. Die Boltzmannsche Formel aus [20], S.173 lautet

$$S = k \log W. \tag{5}$$

In Worten, die Entropie *S* ist proportional dem Logarithmus der *Zustandswahrscheinlichkeit W*. Wir können die Boltzmann-Konstante $k = 1$ setzen, wenn wir den Logarithmus zur Basis 2 nehmen, wie es in der Informationstheorie üblich ist. Dann gilt

$$S = \log W. \tag{6}$$

Was ist nun die Wahrscheinlichkeit eines Strings (1) von der Länge N (in unserem Fall $N = 22$)? Jede Ziffer oder "Bit" 0 oder 1 habe die gleiche *Wahrscheinlichkeit* $w = ½$, entsprechend dem Wurf einer symmetrischen Münze. Eine wichtige Voraussetzung ist auch, dass jede Ziffer statistisch unabhängig von allen anderen Ziffern ist. Dann ist die Gesamt-Wahrscheinlichkeit W gleich dem N-fachen Produkt aller Einzel-Wahrscheinlichkeiten w :

$$W = w^N = (\tfrac{1}{2})^N = 2^{-N} . \tag{7}$$

Die Entropie S ist also nach (6) der Exponent:

$$S = - N. \tag{8}$$

Nun gilt aber nach Brillouin ([20], S. 175)

$$\text{Information} = \text{Negentropie} = \text{negative Entropie} = + N . \tag{9}$$

Die algorithmische Komplexität ist aber nach Chaitin der *Informationsgehalt,* also die Information N, im Einklang mit (2). (Hier habe ich etwas „gezaubert", man hat Gleichheit nur für Strings mit großen N, oder $N \rightarrow \infty$, aber dann kommt man intuitiv zur richtigen Definition; mehr in [3], S. 90 ff.) Die Formel (7) werden wir noch später verwenden.

3.2 Binäre Folgen, nicht-berechenbare Zahlen und Turings Halteproblem

Wir brauchen von der Mathematik nur die ganzen Zahlen 0, 1, 2, 3,… Der Computer kommt eigentlich mit nur 0 und 1 aus (die schon verwendeten binären Folgen). Der Begriff „Bit" für 0 bzw.1 ist aus der Rechentechnik wohlbekannt.

Mit Turing (1936) machen wir so metamathematische Gedankenexperimente, keine wirklichen Computerrechnungen! Aber mit einem idealisierten „Turing-Computer" kann man etwas machen, was schon Kurt Gödel auf rein logisch-mathematischem Weg einige Jahre vorher (1931) gezeigt hatte, man kann die Unvollständigkeit der Mathematik beweisen [22]! Heute ist der Turing-Computer das wichtigste (gedankliche!) Werkzeug der Metamathematik. Siehe **WP**, "Turing machine" (**WP** bedeutet „Wikipedia-Artikel im Internet").

Zur Terminologie

In Abschnitt 3.1 haben wir schon *binäre Folgen* kennengelernt, z.B.

0101001101001101111100…

Sie bestehen also aus 0 und 1.

Wir können daraus eine *binäre Zahl* machen, indem wir als Komma eine „0." vorsetzen:

0.0101001101001101111 00 ...

Der „ganzzahlige Teil" (*integer part*) ist hier immer konventionell gleich Null gesetzt. Dem Computergebrauch folgend verwenden wir als „Komma" einen Punkt (.) statt (,). (Bei Binärzahlen kann man nicht gut die Bezeichnung „Dezimalpunkt" verwenden, aber die Bedeutung ist die gleiche wie bei Dezimalzahlen.) Es gibt endliche und unendliche Binärzahlen, wie auch bei Dezimalzahlen.

Endlichen Binärzahlen entsprechen endliche Binärfolgen, auch *Strings* genannt:

0.01010011 entspricht dem String 01010011. Die oben erstgenannte Zahl 0.010100110100110111100… und die entsprechende Folge 010100110100110111100… sind unendlich.

Berechenbarkeit

Sehr vereinfacht kann man den Turing-Computer als eine Maschine betrachten, die mittels eines Programms p_1 eine binäre Folge als Output liefert:

Turing: $p_1 \rightarrow$ 1 1 0 0 1 0 1……….

Links steht das Programm (einschließlich numerischer Eingabedaten) als *endlicher* Binärstring, rechts ist der Output des Computers als *unendliche* Binärfolge. Turing (1936) hatte noch keinen Computer im heutigen Sinn, wir können aber genauso gut einen heutigen Großrechner und eine moderne Programmsprache wie MATHEMATICA® nehmen, die ja auch letztlich intern im Binärcode arbeitet (Fehler bei der Rechnung werden definitionsgemäß durch Idealisierung vernachlässigt.) Das Zeichen „→" entspricht der im vorigen Abschnitt verwendeten Kolmogorow-Funktion *F* nach Formel (2).

Wir folgen Chaitin [5], S. 8 ff. Wir machen eine Auflistung der Ergebnisse aller unendlich vielen Programme unserer Programmsprache, $p_1, p_2, p_3, \ldots$, und ihrer Ergebnisse im Binärcode; diese Output-Folgen („Strings") werden allgemein unendlich sein.

p_1 1 1 0 1 0 0 1………

p_2 0 1 0 0 0 1 1 ……..

p_3 1 0 0 1 0 1 0 …….. (A)

p_4 1 0 1 0 1 0 0………

Wir haben hier nur Programme aufgelistet, die nach endlicher (und sei es noch so langer) Zeit *anhalten* und damit eine Zahl berechnen. Es gibt auch

Programme, die *nicht halten* und z.B. eine unendliche Schleife ergeben. Solche nicht-haltende Programme haben wir zunächst unbekümmert von der Auflistung ausgeschlossen (ohne uns um das als Nächstes zu besprechende „Halteproblem" zu sorgen).

Durch „Diagonalisierung" kann man eine mögliche unendliche Folge von Binärzahlen erzeugen:

$$d \quad 0\,0\,1\,0\,1\,1\,0\ldots\ldots \tag{B}$$

Dieses Verfahren besteht darin, dass man von der ersten Output-Folge das *erste Glied* nimmt (in unserem Fall also 1) und es durch seinen Komplementärwert (in unserem Fall 0) ersetzt. Von der zweiten Folge nimmt man das *zweite Glied* (in unserem Fall wieder 1) und ersetzt es durch seinen Komplementärwert (in unserem Fall 0). Von der dritten Folge nimmt man das *dritte Glied* (in unserem Fall 0) und ersetzt es durch seinen Komplementärwert (in unserem Fall 1) usw. Auf diese Weise erhält man alle (endlich oder abzählbar unendlichen) Glieder der Folge (B).

Nun sieht man leicht, dass sich die so erhaltene Folge (B) sich von allen Folgen $p_1, p_2, p_3,\ldots$oben unterscheidet: von der ersten Folge durch das erste Glied, von der zweiten Folge durch das zweite Glied, von der dritten Folge durch das dritte Glied usw. Die Folge (B) unterscheidet sich also von *allen* Folgen $p_1, p_2, p_3,\ldots$, im Widerspruch zur Annahme, dass der Raum der Folgen p_i *alle möglichen solcher binären Folgen enthält.*

(Der Mathematiker weiß, dass ein solches Verfahren der Diagonalisierung von Cantor zum Beweis verwendet wurde, dass die Menge der reellen Zahlen *wesentlich* größer ist als die Menge der rationalen Zahlen.)

Das heißt, die Programme p_i ergeben alle *berechenbaren* Zahlen. Die Zahl d wird durch keines dieser Programme berechnet; sie ist *nicht-berechenbar*. Die berechenbaren Zahlen bilden sogar eine Minderheit, ebenso wie die rationalen Zahlen eine Minderheit unter den reellen Zahlen sind.

Halteproblem von Turing

In die obige Aufzählung haben wir nur die „haltenden" Programme aufgenommen. Nun fügen wir alle weggenommenen nicht-haltenden Programme wieder ein und fragen, wie man in dieser „neuen" Aufzählung aller haltenden *und* nicht haltenden Programme die *haltenden* erkennen kann, bevor wir alle Programme laufen lassen. Turing zeigte, dass das nicht möglich ist: Das *Halteproblem* ist unlösbar! Siehe zum Beispiel **WP** "Halting problem".

Zunächst einmal Beispiele: das Programm-Fragment

```
while True: continue
```

hält nie an, es ergibt eine unendliche Schleife, zum Schrecken des Programmier-Anfängers, während das Programm-Fragment

```
Print „Hello, World!"
```

sehr schnell anhält und den Satz „Hello, World!“ ausdruckt.

Ein noch einfacheres Beispiel. Das Programm für

0 / 1 = 0

hält an und führt zu einem vernünftigen Ergebnis, nicht aber das Programm für

1 / 0 = ???,

das zu einer unendlichen Schleife führt. Der Computer hält nicht an, man muss ihn mit Gewalt abschalten. Es sei denn, der Programmierer beugt dem durch einen Kunstgriff vor. Bei einem komplizierten Programm aber kann man von vornherein nicht wissen, ob ein solcher Fall vorliegt!

Das also ist das *Halteproblem von Turing***.** Im genannten **WP**-Artikel gibt es einen eleganten Beweis („Sketch of proof“), dass es nicht möglich ist, das Halteproblem zu lösen.

Die haltenden Programme bilden eine Minderheit; die meisten Programme (als beliebige Strings im Binärcode) halten nicht, ebenso wie wir oben gesehen haben, dass die meisten Zahlen nicht-berechenbar sind.

3.3 Unvollständigkeit nach Gödel

Wir werden nun wieder stark vereinfachen. Nehmen wir an, dass wir nun als Input endlich viele Input-Strings in Binärcode P_i haben (i = 1,2,3….N, N groß), welche als Axiome die Mathematik vollständig beschreiben sollen. Das wäre der Fall des großen Projekts von Hilbert, der vollständigen Axiomatisierung der Mathematik:

Hilbertsche Mathematik: ***Axiome → Deduktion → Theoreme***

oder INPUT (P_i) → OUTPUT (p_j) (j = 1, 2, 3…)

Hier kann man zeigen, dass wir damit niemals alle Theoreme der Mathematik bekommen können. Es liegt also *Unvollständigkeit* vor. Das ist das Wesen des großen *Theorems von Gödel* (1931), dessen exakter Beweis natürlich sehr kompliziert ist [22].

Ein Beispiel für ein bisher ungelöstes Problem könnte die Riemannsche Vermutung in der Theorie der Primzahlen sein. (Als anderes Beispiel wurde in [20] 1995 der berühmte Große Fermatsche Satz genannt, der überraschenderweise inzwischen bewiesen wurde und der daher hier nun wegfällt.)

Das Halteproblem von Turing und das Unvollständigkeitstheorem von Gödel sind weitgehend äquivalent und gehören zu den wichtigsten Ergebnissen der Metamathematik. Die vielleicht einfachste und leicht lesbare, mathematisch strenge Darstellung dieser Problematik ist der schöne Artikel „*What is a computation*?" von Martin Davis in [28], S. 241-267, der auch Chaitins Gedanken über Komplexität und Zufallszahlen in [4] darstellt. Er zeigt auch an konkreten Beispielen, wie man jeden Input P_i als endlichen Binärstring schreiben kann.

3.4 Platons Binärdialektik und Fichtes Wissenschaftslehre nach Speiser

Die binäre Basis (1, 0) spielt als (Eins, das Andere) eine grundlegende Rolle in Platons Dialog „Parmenides" und wird dort kunstvoll verarbeitet [27]. Es zeigt sich, dass Platons Gedanken direkt zu Fichte (Wissenschaftslehre 1804) und Hegel führen. Siehe auch [20], S. 220-225.

Der Physiker und Philosoph Hans-Jürgen Treder, ein bedeutendes Mitglied der Leibniz-Sozietät, liebte Speisers Buch sehr. Der Verfasser dankt Herrn Treder auch an dieser Stelle für interessante Gespräche über Platon und Speiser.

4. Die Chaitin-Konstante

In der gleichen beschreibenden Weise versuchen wir, die Chaitin-Konstante darzustellen, die man vielfach als das dritte wichtige Ergebnis (nach Gödel und Turing) der Metamathematik ansieht. Im Übrigen hängt sie eng mit dem Halteproblem zusammen.

Chaitins Konstante oder Haltewahrscheinlichkeit **Ω** (Omega) ist definiert durch

$$\Omega_F = \sum_{p \in P_F} 2^{-|p|} \tag{10}$$

Hier ist P_F der Definitionsbereich einer („präfix-freien universell berechenbaren") Funktion F, siehe in Wikipedia (**WP**) den ausgezeichneten Artikel "Chaitin's constant"; der String p ist ein Element von P_F (P_F ist ein Unterraum des in Abschn. 5.1 zu erwähnenden *Binärraums* und $|p|$ ist die Länge des Strings p). Die Verwandtschaft dieser Formel mit (7) ist unverkennbar. Der Nichtmathematiker möge sich ruhig schaudernd abwenden, aber schön ist die Formel doch! Vor kurzer Zeit ist auf Deutsch das hervorragende Buch [16] erschienen, das in diesem Aufsatz kaum mehr verwendet werden konnte, ihn aber bis einschließlich dieses Kapitels 4 mathematisch sehr gut ergänzt.

Chaitins Konstante spielt eine wichtige Rolle für die Untersuchung von Entscheidungsproblemen wie Goldbachs Vermutung oder Riemanns Vermutung in seiner Theorie der Primzahlenverteilung (bisher ungelöst); siehe [20], S. 209. Chaitins Konstante liegt gewissenmaßen an der Grenze zwischen Berechenbarkeit und Nicht-Berechenbarkeit im Sinne von Abschnitt 3.2.

Im Weiteren folgen wir Chaitin [7]. Wir nehmen an, dass unser Computer *alle* Input-Strings (d.h. Programme) von der Form P_i der Reihe nach abarbeitet, die einer gewissen Programmsprache, z.B. MATHEMATICA®, entsprechen (wir arbeiten prinzipiell immer im Binärcode). Wie wir gesehen haben, ergeben nicht alle Programme ein „vernünftiges" Ergebnis; wir haben ein Halteproblem. Omega gibt gerade die Wahrscheinlichkeit, dass das betrachtete Programm doch anhält. (Der numerische Wert für Omega hat für jede Programmsprache einen wohldefinierten Wert, der aber nicht einfach zu berechnen ist.) Ein ganz einfaches Beispiel von Chaitin [7] soll das zeigen. Nehmen wir an, unsere Programmsprache bestehe nur aus drei Programmen, die anhalten, nämlich im Binärcode die Strings 110, 11100 und 11110. Ihre Länge beträgt 3, 5 und 5. Die obige Formel für Omega gibt dann mit der Gleichwahrscheinlichkeit ½ für jedes Bit

$$\Omega_F = 1/2^3 + 1/2^5 + 1/2^5 = 1/8 + 1/32 + 1/32 = 0.001 + 0.00001 + 0.00001$$
$$= 0.00110$$

Im Binärcode also eine wohldefinierte Zahl (so einfach ist es nur in diesem höchst vereinfachten Beispiel!). Sonst ist Ω eine ebenso wohldefinierte positive Zahl ≤ 1 (als Wahrscheinlichkeit!), also z.B.

$$\Omega = \Omega_F = 0.00111100011\ldots\ldots$$

wie wir gleich sehen werden. Die *Wahrscheinlichkeit* oder der *Zufall* kommt durch die der Definition zu Grunde liegende Gleichwahrscheinlichkeit ½ zum Ausdruck, welche jedem Bit 0 oder 1 „demokratisch" den gleichen Wert zuteilt, entsprechend dem zufälligen Ergebnis beim Wurf einer Münze (empirisch, aber idealisiert!). Also wiederum, wie im Kapitel 2, können wir vom Zusammenwirken von Gesetz und (idealisiertem) Zufall sprechen.

Obwohl wir gesehen haben, dass Ω nicht direkt berechenbar ist, haben Calude et al. [2] Ω als *berechenbar aufzählbar* (*computably enumerable*) berechnet, d.h. als Grenzwert einer berechenbaren, monoton wachsenden konvergenten Folge von rationalen Zahlen. Für das von ihnen benützte Computerprogramm fanden sie

$$\Omega = 0.\,000000100000001000001100010000110100011111100101\,11011101000010000\ldots \qquad (11)$$

Die angegebenen 64 Bits (Binärziffern) sind exakt, aber die Folge geht noch weiter. Ω ist relativ klein gegenüber 1, was auch die Tatsache ausdrückt, dass statistisch die meisten Programme nicht anhalten.

Argumente von Ebeling. Nun können wir wieder auf die wichtigen Argumente von Ebeling et al. (Abschn. 3.1) zurückkommen. Die asymptotische Formel (9) gibt für die algebraische Komplexität oder den Informationsgehalt K den Ausdruck

$$K = N \qquad (12)$$

für lange Binärfolgen der Länge N. Für unendliche Binärfolgen, die zufälligen (*random*) Charakter haben, gilt daher

$$K \rightarrow \text{unendlich} \quad \text{oder schlichter } K = \infty\,, \qquad (13)$$

was aber nach Ebeling nicht einleuchtet (s. o): diese Folgen sind in keiner Weise strukturiert. Es scheint, dass dieser Ausdruck auch für die Chaitin-Konstante Ω (von primitiven Fällen wie dem obigen Beispiel abgesehen) gilt, die eine unendliche *zufällige* Binärfolge ist, obwohl diese nicht berechenbar, sondern nur *berechenbar aufzählbar* ist. Chaitin hat in der Tat bewiesen, dass Omega „irreducibly complex with *infinite* algorithmic information content" ist („irreducible" in der Computer-Sprache heißt: „nicht mit einem Zip-Programm komprimierbar"). Also ist das Ergebnis weiters nicht verwunderlich, Omega *muss* sogar die Komplexität Unendlich haben, um die unendliche Komplexität der Biologie mathematisch beschreiben zu können, siehe unten. Also für *eine* zufällige Binärfolge, nämlich die von Ω, ist eine unendliche Komplexität doch sinnvoll! In der Tat beschreibt diese Zahl die Struktur der betrachteten Programmsprache. (Gut funktioniert auch der Fall geringer Komplexität, etwa beim Problem des homogenen *Sandhaufens* von Abschnitt 2. Das eindimensionale Analogon ist eine homogene Folge wie 0101010101010101010101010…., die durch ein ganz kurzes Computerprogramm erhalten werden kann und daher sehr geringe Komplexität besitzt. Auch für berechenbare Zahlen wie Pi = 3.142159…. ergibt der Komplexitätsbegriff von Kolmogorow-Chaitin (2) einen vernünftigen Wert.)

4.1 Chaitin und Leibniz

Was hätte wohl Gottfried W. Leibniz zu all dem gesagt? Er wäre sicher stolz gewesen, dass seine binären Zahlen und die gigantische Entwicklung des letztlich auf ihn zurückgehenden Computers (Pascal hatte nur eine Additionsmaschine) einen solchen Aufschwung erhielten; siehe Leibniz, *Explication de l'arithmétique binaire,* 1703, S. 291-293 in Umberto Eco [14]. Er hatte auch eine Idee von Komplexität („*fort composée*") und hätte auch die Zahl **Ω** als platonische mathematische Größe wohl sehr bewundert. Schwierigkeiten hätte er vielleicht mit der Deutung ihrer Binärziffern als zufällige („*random*") Variable gehabt nach seinem *Satz vom zureichenden Grund* (siehe z. B. den Wikipedia-Artikel (**WP**) über G.W. Leibniz):

„Im Sinne des zureichenden Grundes finden wir, dass keine Tatsache als wahr oder existierend gelten kann und keine Aussage als richtig, ohne dass es einen zureichenden Grund [raison suffisante] dafür gibt, dass es so und nicht anders ist, obwohl uns diese Gründe meistens nicht bekannt sein mögen."

(Siehe Leibniz, *Monadologie* und *Theodizee.*) Demnach ist in der Welt nichts zufällig (z.B. A. Einstein: „Gott würfelt nicht").

Die Lösung dieses scheinbaren Widerspruchs ist nach Chaitin vielleicht die dialektische Erkenntnis, dass die Ziffern (Bits) von **Ω** als mathematische Funktionen von ½ eigentlich doch nicht zufällig sind und daher sehr wohl zur platonischen Ideenwelt der Mathematik gehören und folglich *notwendige* Wahrheiten sind**,** obwohl die ½ als Ausdruck von „Gleichwahrscheinlichkeit" doch „ursprünglich" einen *zufälligen* (random) Charakter hat. Wir haben also wieder das für die Komplexitätstheorie so charakteristische Zusammenwirken (Synthese) von Gesetz und Zufall (s.o.). Damit wäre **Ω** ein Beispiel (vielleicht das einzige) einer exakt definierten Zufallszahl (*random variable*), die rein aus der Mathematik, nämlich aus ½, ohne empirische Beobachtung, erhalten werden kann. (Die „Zufallsgeneratoren" erzeugen nur pseudo-zufällige Zahlen!)

In dieser Synthese erinnert Omega an die „coincidentia oppositorum" (Zusammenfallen von Gegensätzen) von Nicolaus Cusanus, also hier von Zufall und mathematischem Gesetz. **Ω** hat für manche Menschen eine geradezu mystische Anziehung.

Von bedeutendem philosophischem Interesse in Bezug auf die tief liegenden Einwände von Ebeling et.al. in Abschnitt 3.1 ist also die Tatsache, dass **Ω** zwar eine Art Zufallszahl ist, aber gleichzeitig *maximale Komplexität* ∞ besitzt, und dass es zwar *nicht* direkt *berechenbar*, aber nach Abschnitt 4 sehr wohl *berechenbar aufzählbar* (*computably enumerable*) ist und einen wohldefinierten Wert hat, der in diesem Sinne auch Bit für Bit numerisch berechnet werden kann. Die algebraische Komplexitätstheorie und Omega spielen gerade für Chaitins Modell der „Metabiologie“ eine wesentliche Rolle, wie wir sehen werden.

In [10], Kapitel 8 („How real are the real numbers?“, S.55-75) gibt Chaitin eine ausführliche Beschreibung der Komplexität der reellen Zahlen. In [8] finden wir einen Überblick über dieses Problem im Zusammenhang mit Leibniz.

Abschließend zitieren wir noch Hoffmann [16], S. 313: Er nennt Omega

„ …eine wahrhaft wundersame Zahl, deren Entdeckung zu den Sternstunden der modernen mathematischen Logik zählt.“

4.2 Die Leibniz- / Chaitin-Medaille

Wir folgen H. Zenil [31], „Leibniz medallion comes to life after 300 years in celebration of Greg Chaitin’s career.”

Am 2. Januar 1697 schickte Leibniz in einem Brief an Rudolph August, Herzog von Braunschweig-Wolfenbüttel, den Vorschlag für eine Gedenkmedaille aus Silber. Der Entwurf sollte die platonische Analogie zwischen „der Erschaffung des Alls aus dem Nichts“ und der Tatsache, dass „alle Zahlen aus 0 und 1 gebildet werden können“ darstellen. Die Rückseite sollte den Herzog Rudolph August zeigen. Die Medaille wurde anscheinend nie geprägt [31]. Der Artikel [15] zeigt auf S. 31-37 den Brief und die Medaille.

Auf Anregung von Stephen Wolfram, dem Schöpfer des bekannten Programm-Systems MATHEMATICA®, arbeitete eine Gruppe von Wolfram Research (Champaign, Illinois, USA) einen Entwurf für die Rückseite einer Leibniz-Medaille anlässlich des 60. Geburtstags von Gregory Chaitin aus. Natürlich sollte im Zentrum die Chaitin-Zahl Ω stehen. Wie im vorhergehenden Abschnitt gezeigt, hatten Calude et al. einen Zahlenwert für Ω berechnet, der in der binären Standard-Formulierung folgendermaßen lautet (die ersten 40 Ziffern werden angegeben):

$$0010000001000010100111011100001111 1010. \tag{14}$$

Der Text der Rückseite ist wie bei Leibniz im klassischen Latein und besagt:

„Alles kann in Einem (UNO) zusammengefasst werden, was selbst nicht direkt erreicht werden kann“ (das UNO wird UNΩ geschrieben, in Anspielung auf die Chaitin-Konstante) und

„Die Wahrheiten der Mathematik zeigen sich als zufällig“
„Zu Ehren von Gregorius Chaitin 2007“

Die Leibniz-Chaitin-Medaille. Links: Entwurf von Leibniz; rechts: Entwurf von Wolfram. (Beide Bilder aus Chaitins Home Page mit freundlicher Genehmigung.)

Die Vorderseite bildet die obige Leibniz-Medaille. Die Leibniz-Chaitin-Medaille wurde von Wolfram am 15. Juli 2007 feierlich vorgestellt. Das Silber-Original wurde am 2. November 2007 an Chaitin überreicht.

5. Anwendung auf biologische Evolution

5.1 Genetischer Code und Binärcode

Wesentlich in der Biologie ist die genetische Struktur, die sowohl das einzelne Lebewesen als auch die biologische Evolution bestimmt. Die genetische Struktur wird durch den *genetischen Code* definiert, welcher aus nur 4 Elementen A, G, C, T besteht, die Abkürzungen sind (A: Adenin, G: Guanin, C: Cytosin, T: Thymin, deren Namen man sich natürlich nicht zu merken braucht); cf. [20], S. 175–177. Das „genetische Skelett“ jeder Zelle ist eine komplizierte „Doppelhelix“, die nur aus diesen Elementen A, G, C, T besteht und für alle Zellen eines Individuums und sogar jeder biologischen Art identisch und daher charakteristisch ist; er heißt auch genetischer Code oder Genom. Die Elemente A, G, C, T sind aber für *alle* Lebewesen gleich, sei es ein Bakterium, ein Regenwurm, eine Rose oder ein Mensch. Ein überzeugendes Argument für die Evolution!

Es ist auch klar, dass die Lebewesen umso höher stehen und definitionsgemäß umso so komplexer sind, je umfangreicher ihr Genom ist. Das Analogon des genetischen Codes (A, G, C, T), durch die DNA repräsentiert (Kein

Mensch spricht heute noch von DeoxyriboNucleicAcid!), ist der oben ausführlich besprochene binäre Code (1.0). Dadurch ist es möglich, die Komplexität durch die Länge der binären Folge zu definieren, was zu einer exakten Definition der Komplexität über die „algorithmische Informationstheorie" führt, wie wir in Abschnitt 3.1 gesehen haben.

Auf diese Weise kommen wir zu einer diskreten Struktur, die mathematisch leichter zu behandeln ist und gewissermaßen „natürlich" ist, denn lineare Strings vertragen sich gut mit der linearen Zeit.

Chaitin schlägt daher folgende Reduktion von der Biologie auf die Mathematik vor:

Lebewesen → DNA (Basen A, G, C, T) → binäre Folgen (Elemente 1, 0).(15)

(Die Elemente des DNA-Codes bezeichnet man gerne als *Basen*, zusammen bilden sie das „Alphabet" des Codes; die molekulare Struktur DNA ist also sozusagen die universelle Programmsprache der Biologie.)

Diese vereinfachende Reduktion ist keineswegs selbstverständlich oder auch nur allgemein üblich, sie ermöglicht aber eine exakte mathematische Behandlung, von welcher der Mathematiker Chaitin ein Verständnis der wesentlichen Vorgänge der biologischen Evolution erwartet.

Für Chaitin grundlegend sind also die Analogien zwischen Biologie und Theorie der Informationsverarbeitung. Das bedeutet nicht, dass ein Lebewesen als Rechenmaschine oder gar als etwas Mechanisches angesehen wird; dafür ist das Lebewesen etwas viel zu Komplexes, was auch in einem sehr tiefen Sinn Kreativität impliziert. Aber die strukturellen Analogien zwischen biologischer Evolution und "evolution of software" sind frappant, nur dass eben die letztere viel einfacher und daher viel leichter zu handhaben und zu einer allgemeinen, vielleicht sogar exakt axiomatisierbaren, mathematischen Theorie („general toy model") führen sollte.

Es bedeutet auch nicht, dass die Schönheit einer Rose oder das Wesen einer Mozart-Symphonie durch ein Computer-Programm beschreibbar oder gar erzeugbar sind. Wohl aber gibt es in der „Mathematik der Mathematik", der *Metamathematik*, Phänomene, die als *Kreativität* interpretiert werden können, cf. auch Penrose [23], S. 416, 423. Die Kreativität (Penrose spricht von *nicht-algorithmischem Denken*) ist also eine typisch metamathematische (Gödelsche) Struktur! Hängt die Zahl Ω mit Kreativität zusammen?

Der Vorteil unseres Spielzeugmodells liegt nun darin, dass es in der binären (Meta-) Mathematik eine präzise und für das binäre Modell sinnvolle Definition durch die von Solomonoff, Kolmogorow und Chaitin entwickelte algorithmische Informationstheorie (AIT) besteht, s. z.B. [3].

Die drei grundlegenden Arbeiten von Chaitin, auf die ich mich in diesem Vortrag besonders beziehe, sind (alle auf Gregory Chaitins Homepage):

- [9] *Mathematics, Biology and Metabiology*, aus dem wir die eingangs angeführten „Aufgaben und Ziele“ haben;
- [10] *Metabiology: Life as evolving software* (PDF, 156 S.). Ein umfassendes Sammelwerk von 7 Arbeiten, das quer durch alle Gedanken Chaitins von Metamathematik zur Metabiologie führt. Es kann nur durch eine gute Kenntnis der *Algorithmischen Informationstheorie* (AIT) völlig verstanden werden, liefert aber auch dem Nichtfachmann interessante Gedanken über Leibniz bis Umberto Eco. Wir werden vor allem solche informellen Gedanken skizzieren.
- [11] *A mathematical theory of evolution and biological creativity*, ein Vortrag, den Chaitin am 10.1.2011 am renommierten Santa Fe Institute of Complexity hielt und der seine neuesten höchst originellen Gedanken und Ergebnisse enthält.

Ein soeben im Druck befindliches Buch von Chaitin [12] gibt einen Überblick über seine biologischen Arbeiten mit hochinteressanten Ausblicken.

5.2 Evolution und Komplexität

Die Komplexität der Biologie ist in der Tat sehr hoch. Wir zitieren wieder Chaitin [10]. Auf S. 75 schreibt er: „Die molokulare Biologie ist ein sehr kompliziertes Gebiet. Eine individuelle Zelle ist [so kompliziert] wie eine Stadt. Jeder von uns hat 3×10^9 Basen in der DNA, also insgesamt 6×10^9 Bits. Es gibt keine einfache mathematische Gleichung für einen bestimmten Menschen. Die Biologie ist die Domäne des Komplexen.

Wie verhält sich die Mathematik zu diesen beiden Gebieten, Physik und Biologie? Normalerweise werden Sie denken, dass die reine Mathematik näher zur Physik liege, weil die beiden Gebiete (seit Leibniz und Newton) sich miteinander entwickelt haben, miteinander aufgewachsen sind. Aber was die Bits der Chaitin-Zahl Ω zeigen, ist, dass in einem ganz bestimmten Sinn die reine Mathematik näher zur Biologie liegt, weil nämlich die reine Mathematik beweisbar *unendliche irreduzierbare Komplexität* enthält. Die Mathematik ist sogar noch „schlimmer“ daran als die Biologie, welche sehr hohe aber *endliche* Komplexität besitzt. Das menschliche Genom ist 6×10^9 Bits, was viel, aber noch endlich viel ist. Die reine Mathematik jedoch enthält die unendlich vielen Bits von Ω !“

Ein anderes markantes Zitat von [9] sagt das Gleiche kürzer aus: “The halting probability Omega has infinite irreducible complexity and is the DNA of pure math, and shows that pure math is infinitely complex and is in this sense closer to biology than to theoretical physics”. (Diese Analogie zwischen Ω und DNA ist natürlich sinngemäß und nicht wörtlich zu nehmen.) Tatsächlich aber spielt Ω eine grundlegende Rolle in Chaitins Theorie der Evolution, wie er in dem erscheinenden Buch [12] eingehend zeigen wird.

Die Komplexität der durch Evolution entstandenen biologischen Welt, von Einzellern bis zu Pflanzen und Tieren, ist natürlich noch viel größer als die eines einzelnen Lebewesens, und daher ist die *Evolution das nächstliegende Gebiet für Chaitins Modell.*

Dieses mathematische “toy model“ nennt Chaitin *Metabiologie*, weil seine mathematische Grundlage die Metamathematik ist, eine mathematische Disziplin seit Gödel und Hilbert. Kurz können wir vergleichen:

Physik ist die Domäne des Einfachen, der kontinuierlichen Mathematik, der Differentialgleichungen;

Biologie ist die Domäne des Komplexen, der diskreten Metamathematik, der algebraischen Informationstheorie, der Chaitin-Zahl Ω.

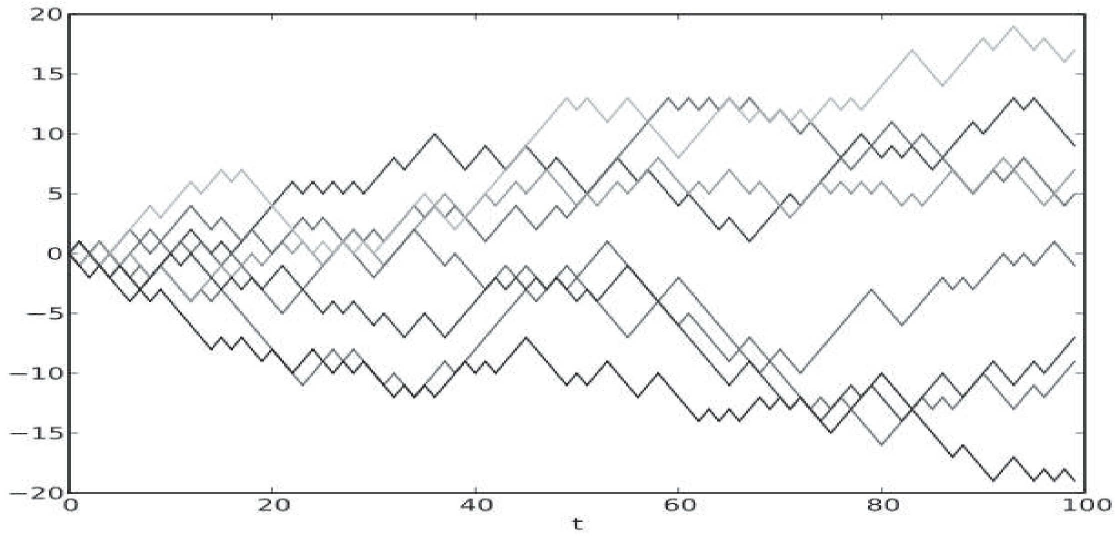
Der historische Verzweigungspunkt von Physik und (Meta-)Biologie liegt bei *Leibniz***,** dem Schöpfer der Infinitesimalrechnung (für die klassische Mathematik) *und* der Binärzahlen (als Grundlage für die Metamathematik).

(*Nur für Mathematiker*: Leibniz‘ Idee von der „Ableitung“ *dy/dx* als Quotienten zweier „unendlich kleinen“ Differentiale *dx* und *dy*, die sehr praktisch ist und auch von Physikern fast immer verwendet wurde, aber bis vor kurzer Zeit bei den Mathematikern als nicht streng verpönt war, wurde vor einigen Jahrzehnten mit Hilfe der mathematischen Logik auf einwandfreie Weise exakt begründet; siehe z.B. D. Laugwitz, *Infinitesimalkalkül*, B-I-Wissenschaftsverlag 1978. Damit lassen sich die Physik als Mathematik des unendlich Kleinen und (Meta-)Biologie als (Meta-) Mathematik des unendlich Großen elegant gegenüberstellen. Leibniz hat also entscheidend dazu beigetragen, sowohl das Reich des unendlich Kleinen als auch das Reich des unendlich Großen in der Mathematik zu erforschen!)

5.3 Evolution als Random Walk

Einen aus allen unendlich vielen binären Folgen bestehenden abstrakten Raum könnte man als *Binärraum* bezeichnen. Ein Unterraum dieses Binärraums wurde in Kapitel 4 als P_F eingeführt. Er wird von Chaitin als *software space* bezeichnet.

So kann Evolution angesehen werden als Zufallsbewegung (*random walk*) in einem Raum von binären Folgen (genauer im Unterraum $P_{F)}$. Eine vereinfachte, aber sehr anschauliche Darstellung eines *Random Walk* in einer Dimension zeigt die unten stehende Abbildung aus dem Wikipedia-Artikel „Random Walk"; die Zeitachse ist horizontal. (Das Fremdwort *Random Walk* ist auch auf Deutsch in der Fachsprache zum Standard geworden; im Deutschen gibt es keinen besseren Ausdruck.)



Random Walk, aus dem gleichnamigen Wikipedia-Artikel „Random Walk"

Mit Bezug auf dieses stark vereinfachte Bild kann man das folgendermaßen zeigen:

- Jeder *Random Walk* stellt eine Evolutionslinie dar.
- Jede Knickstelle („Ecke") bezeichnet einen biologischen Organismus (eine biologische Art oder Spezies).
- Jeder Schritt auf einer Evolutionslinie von einer Ecke zur nächsten ist vom nachfolgenden Schritt unabhängig. Die Wahrscheinlichkeit eines Schrittes ist ½ . Auf einer Evolutionslinie sei die Ecke B sei von der Ecke A um *N* Schritte entfernt. Wegen der Unabhängigkeit der Schritte („independent increments") beim *Random Walk* ist die Wahrscheinlichkeit der Mutation von A nach B gleich

  $(½)^N = 2^{-N}$.

5.4 Eine mathematische Formulierung der biologischen Fitness

Biologische Fitness bedeutet sozusagen, dass in der Evolution der Fittere überlebt, (vgl. [18], S. 30). Mathematische Fitness, nach Chaitin [10], "A mathematical theory of evolution and biological creativity", bedeutet Folgendes: Eine algebraische Mutation ist ein Computer-Programm, das als Input den ursprünglichen Organismus A erhält und als Output den mutierten komplexeren Organismus B liefert. Wenn das Programm aus N Bits besteht, dann hat die Mutation von A nach B die Wahrscheinlichkeit 2^{-N}, wie wir soeben gesehen haben. (Eventuelle Organismen zwischen A und B kann man in diesem vereinfachten Modell als „virtuell“ überspringen.)

Für den Nichtfachmann zeigt das einfache obige Bild des Random Walks doch eine gewisse Anschaulichkeit, auch wenn man auf die mathematischen Details verzichtet.

Für Mathematiker hat der leichtverständliche und bekannte einfache Random Walk, wie wir ihn hier beschrieben haben, für das Chaitin-Problem natürlich auch nur intuitiv-heuristische Bedeutung. Es ist doch bemerkenswert, das Chaitin die gleiche Fomel 2^{-N} auch für den viel komplizierteren Random Walk im Raum binärer Folgen bekommt.

Wir folgen nun weiter Chaitin [11]. Auf seine unnachahmliche Art sagt er: „Was ist nun unser Maß für die Fitness? Nun, um unsere Organismen dazu zu bringen, sich weiter und weiter zu entwickeln, müssen wir sie mit einem mathematischen Problem herausfordern, das nie perfekt lösbar ist, das ein unendliches Maß an mathematischer Kreativität erfordert. Unsere Organismen sind Mathematiker, die versuchen, besser und besser zu werden und mehr und mehr Mathematik zu verstehen. Welches mathematische Problem sollen wir verwenden, um sie zur Weiterentwicklung zu zwingen?

Das einfachste höchst herausfordernde Problem ist das BB-Problem (Busy Beaver, der fleißige Biber), welches mit Turings berühmtem Halteproblem verwandt ist. Was ist das Busy-Beaver-Problem? Es ist das Problem, äußerst große und immer größere positive ganze Zahlen zu nennen.

Warum ist hierfür Kreativität erforderlich? Nun, nehmen wir an, wir haben eine große Zahl N und wollen eine größere Zahl nennen. Wir gehen von N zu $N + N$, zu N mal N, zu N hoch N, zu N hoch N hoch $N\ldots N$ mal. Um also sehr große Zahlen zu nennen, müssen wir Addition, Multiplikation, Exponentiation, Hyperexponentiation usw. erfinden, und das erfordert Kreativität:

Busy-Beaver-Problem: $N{+}N,\ N^2,\ N^N,\ (((\,(N^{N)})^{N})^{N})\ldots.)$ (N mal). (16)

Es gibt im Internet einen schönen Aufsatz vom Spezialisten für Quanten-Computer-Komplexität Scott Aaronson, „Who Can Name the Biggest Number?” [1], welchen ich (G.Ch.) sehr empfehlen kann, denn er zeigt, was für ein fundamentales Problem das ist. (Siehe auch Busy Beaver **WP** (H.M.).)

Das ist also mein Maß für Fitness. Jeder meiner Software-Organismen berechnet also eine einzige natürliche (positive ganze) Zahl, und je größer diese Zahl ist, desto größer ist die Fitness. Der betrachtete Organismus A hat also eine bestimmte Fitness N, und dann versuchen wir eine zufällige Mutation, entsprechend der von mir oben erklärten Wahrscheinlichkeit 2^{-N}, und wenn der resultierende Organismus B eine größere Zahl berechnet, dann ersetzt er A. Sonst versuchen wir, A nochmals zu mutieren."

Soweit Chaitin in seinem Vortrag am 10.1.2011 am Santa Fe-Institut für Komplexität [11]. Es handelt sich natürlich noch nicht um eine fertige Theorie, sondern um ein seit etwa drei Jahren laufendes Forschungsprogramm, das noch lange nicht abgeschlossen ist. Es liegen jedoch schon hochinteressante Zwischenergebnisse vor, über die wir hier berichtet haben. Der grundlegende mathematische Beweis findet sich im Anhang 2 seines Buchs [12], das gerade im Druck ist. Dieser Beweis verwendet in entscheidender Weise die Chaitin-Zahl Ω.

Literatur

[1] S. Aaronson , Who Can Name the Bigger Number? http://www.scottaaronson.com/writings/bignumbers.html

[2] C. S. Calude, M. J. Dinneen und Chi-Kou Shu, Computing a Glimpse of Randomness, Experimental Mathematics, 11(3), p. 369-378, 2002.

[3] G. J. Chaitin, Information, Randomness & Incompleteness – Papers on Algorithmic Information Theory, World Scientific Publ. 1987, Singapore (technisch).

[4] G. J. Chaitin, Randomness and Mathematical Proof, Scientific American 232 (5), S. 47-52,1975.

[5] G. J. Chaitin, The Limits of Mathematics, Springer, Singapore 1998 (gut lesbar).

[6] G. J. Chaitin, META MATH ! −The Quest for Omege , Vintage, New York, 2005 (höchst informativ).

[7] G. J. Chaitin, The Limits of Reason, Scientific American, S. 74–81, März 2006 (sehr informativ).

Auf Chaitins Homepage www.umcs.maine.edu/~chaitin zu finden (gut lesbar):

[8] G. J. Chaitin, Leibniz, Complexity and Incompleteness, 2009a;

[9] G. J. Chaitin, Mathematics, Biology and Metabiology, 2009b;

[10]G. J. Chaitin, Metabiology: Life as evolving software (PDF, 156 S., 7 Kapitel), 2010;

[11]G. J. Chaitin, A mathematical theory of evolution and biological creativity, 2011.

[12]G. J. Chaitin, PROVING DARWIN – Making Biology Mathematical, Pantheon, New York 2012 (im Druck).

[13]W. Ebeling, J. Freund, F. Schweizer, Komplexe Strukturen: Entropie und Information, B.G.Teubner, Leipzig 1998.

[14]U. Eco, Auf der Suche nach der vollkommenen Sprache, DTV 2002.

[15]A. Glaser, History of Binary and Other Nondecimal Numeration, www.eipiphiny.org/books/history-of-binary.pdf

[16]D. W. Hoffmann, Grenzen der Mathematik: Eine Reise durch die Kerngebiete der mathematischen Logik, Spektrum Akademischer Verlag 2011 (soeben erschienen, von G. Chaitin besonders empfohlen).
[17]B. Hofmann-Wellenhof und H. Moritz, Physical Geodesy, Springer, Wien 2005.
[18]B. Moritz und H. Moritz, „Naturgesetze und Evolution“, S. 32, in www.helmut-moritz.at
[19]H. Moritz, Advanced Physical Geodesy, Wichmann, Karlsruhe 1980.
[20]H. Moritz, Science, Mind and the Universe – An Introduction to Natural Philosophy (1995), in www.helmut-moritz.at (Hintergrundinformation, als [20] abgekürzt)
(Russische Übersetzung Moskau 2010.)
[21]H. Moritz, Große Mathematiker und die Geowissenschaften, Vortrag Berlin 2008, deutsch in „Sitzungsberichte der Leibniz-Sozietät der Wissenschaften“, Bd. 104 (2009), S. 115 -130; englisch (2010) in www.helmut-moritz.at
[22]E. Nagel und J. R. Newman, Goedel’s Proof, New York Univ. Press 2001.
[23]R. Penrose, The Emperor’s New Mind, Oxford Univ. Press 1989.
[24]Sammelband „Einfachheit als Wirk-, Erkenntnis und Gestaltungsprinzip“, Sitzungsberichte der Leibniz-Sozietät der Wissenschaften, Bd. 108, 2010
(für den gegenwärtigen Vortrag relevante Literatur aus dem Umfeld der Leibniz-Sozietät).
[25]R. Schimming, Back to Bertalanffy: the System Theoretical Approach to Biology, European Communications in Mathematical and Theoretical Biology, 5, 2003, S 11–15.
[26]H. G. Schuster, Deterministic Chaos, VCH, Weinheim 1988.
[27]A. Speiser, Ein Parmenideskommentar – Studien zur Platonischen Dialektik, Kohler, Stuttgart 1959.
[28]L. A. Steen (ed.), Mathematics Today, Springer, New York 1978.
[29]M. M. Waldrop, Complexity: the Emerging Science at the Edge of Order and Chaos, Simon & Schuster, New York 1992.
[30]H. Zenil, Busy Beaver, http://demonstrations.wolfram.com/BusyBeaver/
[31]H. Zenil: Leibniz medallion comes to life after 300 years in celebration of Greg Chaitin’s career, www.mathrix.org/liquid/archives/the-history-of-the-chaitin-leibniz-medallion

Weitere Literaturhinweise findet man im Text; **WP** bedeutet „Wikipedia-Artikel“ (in der Fassung von 2011).

Dank. Der Verfasser dankt Herrn Gregory Chaitin für anregende Diskussionen, für einen Vorabdruck des Buchs [12] und für das Lesen des vorliegenden Manuskripts. Den Teilnehmern an der anregenden Diskussion nach dem Vortrag an der Leibniz-Sozietät sage ich Dank.