

Ideen zu »Am Anfang war die Information«

Matthias Ansorg

01. Oktober 2003 bis 27. Februar 2004

Zusammenfassung

Ideen zur Vorstellung und Diskussion des Buches »Am Anfang war die Information« von Prof. Dr. Werner Gitt. Ich sage deutlich, dass es eine Ideensammlung ist. Sie mag Hinweise zum Umgang mit der Naturgesetzlichen Informationstheorie von Gitt geben, stellt aber keine fundierte Ausarbeitung dar. Dieses Dokument ist public domain, das heißt ich erhebe keinerlei urheberrechtliche Ansprüche.

1 Gedanken zur Definition von Leben

- »Definieren wir: Leben ist die Eigenschaft einer Maschine, sich selbst reproduzieren zu können (von-Neumann-Maschine). Computerviren sind demnach auch Leben.« Aber: die Bibel sagt: »ohne Geist ist der Leib tot« - ist hier mit »tot« »geistlich tot« gemeint, also mit dem Geist der Heilige Geist? Vgl.: »der Geist ist es, der lebendig macht, das Fleisch nützt gar nichts«. Oder ist es tatsächlich der Geist, der ein Stück Materie zu einem lebendigen Menschen macht?
- »Leben ist die Fähigkeit, selbständig nichtzufällige Strukturen hervorzubringen, die nicht durch Naturgesetze erzwungen werden.« Denn auch Lebewesen, die ihre Fortpflanzungsfähigkeit verloren haben, leben noch. Mit dieser Definition würden aber auch autonome Roboter als Leben bezeichnet!
- »Lebendig ist, sich in einem labilen Gleichgewicht zu halten.« Diese Definition passt auf den klinischen Tod, der von außen durch Reanimation behoben werden kann, sofern das »labile Gleichgewicht« noch nicht vollständig zerstört ist.
- »Leben ist die Fähigkeit, unter bestimmten Bedingungen selbständig nichtzufällige und nicht durch Naturgesetze erzwungene Prozesse durchzuführen.«

2 Ebenen der Information

Shannon betrachtet die »materielle Implementierung der Information« als Zeichenfolge (Ebene der Statistik). Syntax ist in Zeichenfolgen implementiert, führt aber zusätzliche Regeln ein (welche Zeichenfolgen zulässig sind). Semantik ist in korrekter Syntax implementiert, führt aber zusätzliche Regeln ein (welcher Syntax bedeutungstragend ist und was die Bedeutung ist). Pragmatik ist die Reaktion auf korrekte Semantik, jedoch nicht (!) in Semantik implementiert. Apobetik ist die Folge von Handlungen (Pragmatik), jedoch nicht in Pragmatik implementiert.

Was sind diese »Regeln«? Kreative, frei erfundene Zuordnungen nach Art einer Relation: nicht alle Tupel (Wort,Bedeutung) sind vorhanden, sondern nur manche.

Jeder mögliche Kode kann binär repräsentiert werden. Der Binärkode ist der strukturell einfachste Kode und kann damit als künstlicher Universalkode verwendet werden.

Die Ebenen »Statistik« und »Syntax« dienen der Abbildung von Semantik in Materie und umgekehrt. »Reine Semantik« ist immateriell, nicht in Materie implementiert. Man kann sie also in der materiellen Welt nicht finden und kann nicht mit ihr umgehen. Das bedeutet nicht, dass es keine »reine Semantik« gibt (Geistleib).

Statistik und Syntax haben mit Information »an sich« also nichts zu tun, sondern sind Werkzeuge zu Umgang mit Information (besser: ihren Abbildern) im materiellen Bereich. Dabei müssen die durch Naturgesetze gegebenen Randbedingungen eingehalten werden; deshalb werden in der Nachrichtentechnik und Natur verschiedene Kodes verwendet, je nach Anwendungsfall der jeweils am besten geeignete.

Achtung: Es ist wohl doch eine gute Idee, unter »reiner Semantik« nicht die bezeichneten Dinge an sich zu verstehen, sondern die »reine Stellvertretung«. Ohne Codesystem. Diese reine Stellvertretung wird auch von Fotos geleistet. Diese sind jedoch nur zur Stellvertretung bloßer Gegenstände geeignet, denn mangels Syntax und Zeichenvorrat sind sie nur ein sehr elementares Ausdrucksmittel (nicht vollständig und nicht eindeutig decodierbar,

sofern es um mehr als Gegenstände geht; vgl. Gemälde und die Gedanken des Künstlers ...). Die Semantik kann einigermäßen »codelos« nur durch ein semantisches Modell vollständig ausgedrückt werden (etwa in UML). Modelle sind reine Stellvertretung, reine Semantik. Semantik kann materielle wie immaterielle (geistige) Dinge beinhalten (stellvertreten). Der Leib für den Geist ist also nicht ein Interface zur reinen Semantik, sondern zu den geistigen Dingen, von denen wir hier nur über ihre Stellvertretung durch Informationen erfahren. Oder anders: wir könnten davon erfahren, denn wir haben bereits einen Geist; die Schnittstelle ist aber nicht geeignet, solche Dinge direkt zu erkennen, sondern nur (undeutlich) über ihre in Materie formulierte Stellvertretung (die Information).

Zu Pragmatik und Apobetik: es scheint abweichend von Gitt vernünftig, die Information als »materiell formulierte Stellvertretung« zu definieren, also mit den Ebenen Statistik, Syntax, Semantik. Pragmatik und Apobetik sind dagegen bereits mögliche Semantiken, aber keine notwendigen (ein Roman fordert nicht zu einer Handlung auf, und der Autor hat keine besondere Zielsetzung beim Leser damit). Es sind bereits »geistige Größen«, die durch Information stellvertreten werden können, aber nicht müssen. Auch Messdaten haben Statistik, Syntax, Semantik - sie haben Stellvertreterfunktion, aber keine Pragmatik und Apobetik. Solcherart »neue Information« könnte durch Geräte erzeugt werden. Vielleicht ist das der Grund, warum Pragmatik und Apobetik in die Definition aufgenommen wurden, denn nur dann ist das Gesetz gültig, das einen geistigen Urheber fordert?

3 Information als geistige Größe

Gitt, Satz 13: »Ein Codesystem entstammt immer einem geistigen Prozess« - wie wäre es mit vernetzten Computern, die ein Codesystem vereinbaren, indem ein Rechner die zu kodierenden Zeichen im ASCII-Kode oder einem anderen existierenden Kode hinschickt und der Partnerrechner ein zufällig gewähltes Binärmuster zurückschickt, das zukünftig zur Darstellung dient?

Die Codebildung ist ein geistiger (kreativer) Prozess - warum? Weil Codes nicht durch Naturgesetze erzwungen werden. Sondern Codes verwenden Naturgesetze.

Ist vielleicht der Geist (des Menschen oder Gottes) die einzige Entität, die »reine Information« verarbeiten kann? Der menschliche Geist kann vielleicht solange keine reine Information verarbeiten, wie er im materiellen Körper steckt. Alle materialisierte Information, selbst wenn sie autonome materielle Systeme darstellt (Leben) ist darum nichts als ein Ausdruck der »reinen Information« im materiellen Bereich. Dazu sind Kodes notwendig.

Ist doch komisch: Neuronale Netze können Information verarbeiten, verwenden aber keinen Kode (keine abgegrenzte Symbolmenge usw.). Ein Grund, warum das Hirn zur Verarbeitung »reiner Information« ansatzweise fähig ist? Oder sollte man sagen: Kodes sind nur zur Informationsübertragung nötig, nicht zur Informationsdarstellung bzw. -speicherung »innerhalb des eigenen Systems«.

Was ist »Bedeutung«, »Semantik«, synonym »reine Information«, »Information vor ihrer materiellen Implementierung«? Man könnte sagen: Bedeutung ist die Implementierung einer Absicht, Absichten sind beliebige Äußerungen des »freien Willens«.

Wenn Semantik eine Zeichenkette erst zur Information macht ([?, S. 110]), was ist dann die Bedeutung der Tiere? Was will Gott damit sagen? Die Semantik ist das Wesentliche der Information, weil sie invariant ist bei Änderung der Darstellung [?, S. 111].

»Der Informatiker W. Strombach [S12] definierte Information als eine Struktur, die in einem empfangenden System etwas bewirkt.« [?, S. 116] Das empfangende System ist sozusagen der »Prozessor«.

»NGI-6: Es gibt keine neue Information ohne einen intelligenten und mit Willen ausgestatteten Sender.« [?, S. 141]. Kreativität ist, Systeme von Regeln erfinden. Kreativität führt stets zu neuer Information, aber nicht jede neue Information ist kreativ, denn nicht jede neue Information stellt ein System von Regeln dar.

Als »eigentlicher Sender« wird der identifiziert, der wollte, dass etwas ist oder abläuft. Denn es gibt keine Information ohne Willen.

Zu NGI-8: Was ist »verstehen«?

Alternative 1: Verstehen ist ein Prozess, der für jede empfangene Information neu notwendig ist. So ist »Verstehen« in der Umgangssprache belegt. Dann aber ist »Verstehen« kein geistiger Prozess, sondern für geistlose Maschinen möglich. Dass diese Maschinen Produkte eines geistigen Prozesses sind, ist dabei belanglos.

Alternative 2: Verstehen ist der Prozess, für eine begrenzte Menge möglicher Informationen durch Regeln festzulegen, welche Bedeutung, Handlung und Ziel damit verbunden ist. Dieser Prozess ist aus der »Codedefinition« automatisch als Umkehrung durchführbar. Die Codedefinition ist ein geistiger Prozess, da er freien Willen braucht. Damit gibt es auch die Umkehrung nicht ohne Geist. Verstehen wäre nur einmal »allgemein« nötig, etwa wenn das Verständnis in einem Programm formuliert wird, das dann jeder empfangenen Information automatisch die richtige Bedeutung zuordnen kann.

Der Sender muss dafür sorgen, dass ein System zur korrekten Dekodierung zur Verfügung steht. Denn nur der Sender kennt den Code, nur der Sender kann korrekt dekodieren. Nur der Sender versteht, er allein kann die Fähigkeit zum Verständnis vergeben.

Dazu: wenn etwa Leben keine nichtmaterielle Größe enthält, was muss dann gelten? Dann muss das Phänomen Leben allein aus der Materie und den darauf definierten Gesetzmäßigkeiten (den Naturgesetzen) hergeleitet werden können. Die Naturgesetze müssen Leben in Materie mit einer Wahrscheinlichkeit $p \leq 1$ erzwingen. Vgl. [?, S. 68-78] für eine Darstellung, dass es bisher trotz aller Anstrengungen nicht gelungen ist, Leben aus Physik und Chemie herzuleiten. Vgl. [?, S. 80-86].

Definition: semantischer Informationsdichte. Sei a_e die Anzahl der in einer Nachricht explizit enthaltenen Aussagen und a_i die in einer Nachricht enthaltenen impliziten Aussagen. Implizite Aussagen können aus den expliziten Aussagen auf logischem Wege hergeleitet werden, wozu Intelligenz Fähigkeit zum freien Verknüpfen) notwendig ist. Dann ist die semantische Informationsdichte:

$$d_a = 1 + \frac{a_i}{a_e}$$

Am Beispiel der Bibel ist das gut zu beobachten: Nur die allerwenigsten enthaltenen Aussagen sind explizit formuliert, der aller größte Teil kann / muss hergeleitet werden. Damit ist d_a sehr groß.

Möglich als Beispiel für $1 : 10^{2000}$, die Zahl mit der ein selbstreproduzierendes System per Zufall entsteht: ist es vernünftig, Lotto zu spielen? Nein, denn man verliert im statistischen Mittel mehr als man gewinnt. Ist es also vernünftig, an Evolution zu glauben entgegen der Wahrscheinlichkeit $1 : 10^{2000}$? Nein.

Naturgesetze der Physik und Chemie können sowohl zwingende Kausalzusammenhänge sein als auch statistische Zusammenhänge. Ein solcher statistischer Zusammenhang, der aus Naturgesetzen abzuleiten ist, ist die Entstehung von Aminosäuren in einer bestimmten Konzentration bei Miller-Experimenten. Ein ähnlicher statistischer Zusammenhang ist die Konzentration, mit der informationsverarbeitende, selbstreproduzierende Maschinen (von-Neumann-Maschinen wie etwa Pflanzen) von Naturgesetzen erzwungen werden. Sie beträgt $1 : 10^{2000}$. Bei nur 10^{80} Atomen im ganzen Universum bedeutet das $1 : 10^{2000} \Leftrightarrow 0$. Die Naturgesetze über Information sind also ebenfalls statistische Zusammenhänge, jedoch sind die Wahrscheinlichkeiten dabei derart eindeutig, dass sie völlig äquivalent zu zwingenden Kausalzusammenhängen sind.

4 Informatik und Information

Heute: sechs verschiedene Bedeutungen, Information als

- Selbstformung, Selbstbildung: Lernen durch Versuch und Irrtum
- Fremdformung: Bildhauerei, Malerei
- Lehre und Unterweisung
- Weitergabe von Sachverhalten
- Deutung und Erläuterung
- Vorstellung, Wissen, Begriff, Nachricht: »die Vergegenständlichung von Semantik«. Kein Verfahrensbegriff wie die bisherigen. Besonders für die Naturwissenschaft bedeutsam, weil Information so lokalisiert werden kann.

Der Shannonsche Informationsbegriff. Es entstand eine Neudefinition von Nachricht und Information. Shannon vernachlässigte die inhaltliche Bedeutung von Information völlig - und absichtlich.

Darstellen, dass der Shannonsche Informationsbegriff völlig unzureichend ist, weil eben der Inhalt nicht berücksichtigt wird.

Shannon forderte für eine geglückte Übertragung: Ein gemeinsamer Zeichenvorrat muss bei Sender und Empfänger vorhanden sein. Shannons Theorie ist für die Informatik bedeutsam. Jedoch können andere Wissenschaften damit nichts anfangen da ja der Inhalt vernachlässigt wird. Deshalb wurde versucht, den Informationsbegriff neu zu definieren. Dabei kam bis heute keine einheitliche oder allgemein brauchbare Definition heraus.

Gitt sagt, er habe eine solche Definition gefunden, die naturwissenschaftliche Aussagen ermögliche.

Das Sender-Senke-Empfänger-Schema ist eine Erweiterung des Shannonschen Informationsbegriffs.

Der sprachwissenschaftliche Informationsbegriff. Hier wird Information in Syntax, Semantik und Pragmatik zerlegt. Semiotik befasst sich mit dem Prozess, in dem ein Zeichen fungiert.

Definition der Semiose nach Morris 1988. Pragmatik meint »die Konsequenzen von Informationssystemen auf das gesellschaftliche Umfeld«. Nach Morris werden Semantik und Pragmatik getrennt: Pragmatik ist nicht in Semantik implementiert, nicht einmal vom Sender beabsichtigt.

»Struktur der Semantik heißt Pragmatik«

5 Information und Selbstorganisation (Wersig 1971)

Hier wurde von der Evolution der Natur als Tatsache ausgegangen: »Im Laufe der Evolution haben sich Informationsstrukturen gebildet.« Die Informationsstrukturen im Gehirn sind bisher fast völlig unverstanden.

Wechselwirkung: alles hängt mit allem zusammen. Widerspiegelung: eine Wechselwirkung, bei der sich das Innere eines Materiebereichs durch äußere Einflüsse ändert. Dies sei etwas »wie Masse oder Energie« und lege den Grundstein für die Entstehung von Informationsstrukturen. »Die Wirkung der Widerspiegelung tritt als Information auf.«

Widerspiegelungsebenen:

- biologische Widerspiegelungsebenen: ein Geißeltierchen hält sich über Wasser, weil es weiß, wie es auf seine Sinneseindrücke reagieren muss, um sich über Wasser zu halten.
- emotionale Widerspiegelungsebenen:
- geistliche Widerspiegelungsebenen: sich etwas vorstellen
- sprachlich, bildlich: durch Medien
- wissenschaftlich:
- technisch:
- gesellschaftlich:

Widerspiegelung und Selbstorganisation seien die Brücken der sprachwissenschaftlichen Informationsdefinition zu den Naturwissenschaften. Der Text war sehr philosophisch und kompliziert gehalten. Es gibt viele unterschiedliche Modelle, was Information inkl. Semantik sein soll. Das konnte bisher nicht übereinstimmend erfasst werden.

6 Informatik ist Informationswissenschaft

- »Die Produkte der Informatik sind Handlungsanweisungen.«
- Der Gegenstand der Informatik sind die Menschen, die Maschinen und ihre Schnittstellen. Soweit es um Informationsaspekte geht. Informatik beschäftigt sich mit der Serialisierung (also Formalisierung) von geistigen Zusammenhängen.
- »Gegenstand der Informatik ist die Repräsentation von Wissen in Form von Daten und die Reduktion von geistigen Tätigkeiten auf Algorithmen und maschinell implementierbare Prozesse.«
- In der Informatik hat man es immer mit Anwendungen von Informationssystemen zu tun, also mit Menschen und nicht mit Technik allein!