

Wissenschaftliche Arbeit

Informationstheorie und Schriftsprache

von

Sylvia Richter

Matr.-Nr.: 720575

Universität Potsdam
Institut für Informatik
Professur Informationsverarbeitung und Kommunikation

Aufgabenstellung und Betreuung:
Prof. Dr. Helmut Jürgensen

Potsdam
21. Februar 2006

Inhaltsverzeichnis

1	Einleitung	1
2	Information	2
2.1	Gebundene und freie Information	2
2.2	Strukturelle, funktionale und pragmatische Information	3
2.3	Veranschaulichung der Begriffe	4
3	Entropie	5
3.1	Die wichtigsten Eigenschaften	5
3.2	Maximale Entropie	6
3.3	Relative Entropie	7
3.4	Entropie in Sprachen	8
4	Redundanz	10
4.1	Redundanz und ihre Eigenschaften	10
4.2	Redundanz in Sprachen	10
4.2.1	Strukturelle Redundanz	11
4.2.2	Funktionale Redundanz	12
5	Veranschaulichung der pragmatischen Information nach von Weizsäcker	13
6	Zusammenfassung	14

1 Einleitung

Die Informationstheorie, deren Entwicklung vor allem C. E. Shannon im Jahre 1948 voranbrachte, betrachtet die „natürliche Sprache als ein System der Nachrichtenkodierung, wobei verschiedene Eigenschaften dieses Systems durch die Heranziehung statistischer Methoden studiert werden“ [EV73]. Diese Informations- bzw. Nachrichtenübertragung geschieht mittels eines Kommunikationssystems, das mindestens aus den folgenden Komponenten besteht: Quelle, Kanal mit Störung und Empfänger.

Die Arbeit hat zum Ziel, die informationstheoretische Aspekte darzustellen und sie auf die Sprache zu beziehen. Sie soll hauptsächlich als Einführung in die Thematik dienen. Mit dem Begriff Sprache wird im Folgenden stets die Schriftsprache gemeint sein.

Der Begriff Information, der spezifiziert wird in gebundene und freie Information (Kapitel 2.1.) sowie strukturelle, funktionale und pragmatische Information (Kapitel 2.2.), soll in dieser Arbeit erläutert werden. Diese Unterscheidungen werden im Kapitel 2.3. anhand eines Beispiels veranschaulicht.

Im Kapitel 3 wird der Begriff der Entropie mit seinen wichtigsten Eigenschaften (Kapitel 3.1) erläutert und dann speziell auf die maximale (Kapitel 3.2) und relative Entropie (Kapitel 3.3) eingegangen. Diesen Themenkomplex runden Betrachtungen der Entropie in Sprachen ab.

Im Kapitel 4 setzt sich diese Arbeit mit der Redundanz, ihren Eigenschaften und der Redundanz in Sprachen auseinander. Hierbei versuche ich die Begriffe der strukturellen und funktionalen Information auf die Redundanz zu übertragen. Dies ergibt dann die strukturelle und funktionale Redundanz.

Die drei großen Themenkomplexe werden mittels einer Veranschaulichung der pragmatischen Information von C. F. von Weizsäcker abgeschlossen.

2 Information

Das alltagssprachliche Verständnis des Begriffs Information ist herleitbar aus dem Lateinischen *informare* (gestalten, formen, bilden), das im Spätmittelhochdeutschen des 14. Jahrhunderts die erweiterte Bedeutung von unterrichten, durch Unterweisung bilden, befähigen erhalten hat. Um Information zu übermitteln, sind zwei Gesprächspartner, also eine binäre Relation von Nöten. „Eine Information, die niemand teilt, ist keine Information.“ [EFS98]

2.1 Gebundene und freie Information

Nach C. F. von Weizsäcker (1974) sind Masse und Energie Information. T. Stornier (1991) äußert sich ähnlich: „Information existiert. Um zu existieren, muss sie nicht wahrgenommen und nicht verstanden werden. ... Sie braucht keine Bedeutung, um zu existieren. Sie existiert einfach.“ [EFS98] Somit betrachtet Stornier gebundene Information als bereits gegebene Information. Beispielsweise tragen physikalische Strukturen gebundene Information.

Diese gebundene Information kann in freie Information umgewandelt werden. Erst dadurch kann Information mittels Symbolen gespeichert, ausgetauscht und weiterverarbeitet werden. Die freie Information ist, wie bereits erwähnt, eine binäre Relation, die die Beziehung zwischen dem Sender und dem Empfänger ermöglicht. Voraussetzung hierbei ist, dass sowohl Sender als auch Empfänger diese Symbole erzeugen und verstehen können. Die freie Information kann gebundene Information extrahieren, speichern, austauschen und in wiederum gebundene Information weiterverarbeiten.

Nach [EFS98] existiert freie Information nur innerhalb der Welt der Lebewesen. Ferner gehörten Computer zu den Lebewesen, da sie „freie Information austauschen“ [EFS98] und „eindeutig in das System der sozialen Evolution“ [EFS98] eingeordnet werden könnten. Für manchen Leser mag dies skurril klingen, da Computer mit Lebewesen auf eine Ebene gestellt werden und bisher noch nicht nachgewiesen wurde, dass künstliche Intelligenz existiert. Es ist jedoch zu beachten, dass der Inhalt der freien Information nicht wesentlich ist. Selbst die konkrete materielle Bindung auf einem Medium sowie die Codierung sind irrelevant, solange die Kommunikationspartner die gleichen Symbole verstehen.

2.2 Strukturelle, funktionale und pragmatische Information

Der Begriff der strukturellen Information wurde von F. Schweitzer 1977 [EFS98] geprägt. Es ist diejenige Information, die zu einem diskreten Zeitpunkt und an einem bestimmten Ort mit einer materiellen Struktur gegeben ist. Sie repräsentiert als Information erster Art die strukturelle Determiniertheit eines Systems und codiert auf materieller Grundlage Sachverhalte. Sowohl die gebundene als auch die freie Information weisen eine Struktur auf und enthalten deshalb stets strukturelle Information. Folglich repräsentiert die strukturelle Information den syntaktischen Aspekt der Information.

Dem gegenüber steht die funktionale Information. Sie ist die so genannte Information zweiter Art. Diese aktiviert, interpretiert bzw. deutet die Information erster Art. Sie ist kontextabhängig und repräsentiert die von einem Rezipienten aus struktureller Information heraus gelesene und interpretierte Bedeutung der Information. Folglich ist die funktionale Information subjektiv. Gebundene Information kann mittels Symbolen zu freier Information umgewandelt und somit auch ausgetauscht werden. Ein Beispiel für funktionale Information stellt der Algorithmus dar. Er wurde aus struktureller Information gewonnen, indem das zu realisierende System gedeutet wurde. Durch den Übergang von gebundener zu freier Information wurden Symbole verwendet. Ein weiteres Beispiel stellt der Bauplan eines Hauses dar. Er wurde ebenfalls aus struktureller Information gewonnen. Die nun freie Information kann wieder in gebundene Information umgewandelt werden, dadurch, dass das Haus gebaut wird.

Wenn strukturelle Information durch die funktionale Informationsverarbeitung des Empfängers richtig interpretiert wird, so bedeutet dies noch nicht, dass diese Information beim Empfänger auch eine Wirkung auslöst. Diese Umwandlung, wie sie beispielsweise beim Lesen dieses Textes geschieht, passiert selbständig. Sie hängt von dem Algorithmus ab, mit dem Buchstabensequenzen verarbeitet werden. Die pragmatische Information beinhaltet, dass Information zunächst anhand von bereits vorhandenem Wissen verstanden werden muss, um eine entsprechende Wirkung beim Empfänger zu erzielen. Der Zusammenhang zwischen den unterschiedlichen Arten von Information kann wie folgt formuliert werden: „... durch die funktionale Information wird der Übergang von struktureller zu pragmatischer Information ermöglicht (Schweitzer)“ [EFS98]. Nach dieser Sichtweise ist die funktionale Information jene „individuelle Größe des Empfängers“, die E. v. Weizsäcker „als Vorwissen vorausgesetzt hat, während die struk-

turelle Information eine unspezifisch codierte Information ist, die keine Bedeutung mitbringt“ [EFS98].

Die pragmatische Information wird in der Informationstheorie vernachlässigt, da es darum geht, „ob der Empfänger informiert wird oder nicht; ob er auch der Information gemäß handelt, ist ihr grundsätzlich gleichgültig“ [Sei68].

2.3 Veranschaulichung der Begriffe

Werden beispielsweise die Buchstaben des deutschen Alphabets und ihre natürlichen Auftretenswahrscheinlichkeiten betrachtet, so ist die Buchstabenfolge L-O-E-W-E wahrscheinlicher als Y-Y-Y-Y [EFS98]. Jedoch besitzt die Folge L-O-E-W-E einen geringeren Informationsgehalt, was unserem intuitiven Verständnis von Information widerspricht. Dies liegt daran, dass bei der Begriffsbildung von Shannon nur die syntaktische Information, nicht aber die semantische betrachtet wird. Semantisch messen wir dem Wort „Löwe“ eine Bedeutung bei. Erst anhand der pragmatischen Information, also durch das Verstehen auf Grundlage von bereits vorhandenem Wissen, wird eine Wirkung ausgelöst.

Nehmen wir an, wir befänden uns in der Steppe und würden einen Löwen sehen. So wird ein Teil der strukturellen Information in freie Information an uns übertragen. Falls wir noch nie einen Löwen gesehen haben und auch nichts über ihn wissen, würde die Information keine Wirkung bei uns auslösen. Da wir jedoch gebildet sind und funktionale Information über einen Löwen als Vorwissen besitzen, ist es möglich die freie Information in pragmatische Information zu verarbeiten. Dies könnte zur Folge haben, dass eine Wirkung ausgelöst wird. Beispielsweise könnten wir versuchen, weg zu rennen.

3 Entropie

Der Begriff Entropie wurde ursprünglich von Rudolf Clausius 1850 zur Charakterisierung thermodynamischer Prozesse aus der Wärmelehre eingeführt. Um 1880 führten Boltzmann und Planck die Entropie eines Makrozustandes in der statistischen Mechanik als Logarithmus der thermodynamischen Wahrscheinlichkeit ein. Auf diese Definitionen wird im Folgenden nicht näher eingegangen. Falls Interesse über die Beziehungen der statistischen und der informationstheoretischen Entropie bestehen sollte, ist wiederum [EFS98] empfehlenswert.

1948 führten Shannon und Hartley die informationstheoretische Entropie ein, durch die eine quantitative Messung des durchschnittlichen Informationsgehaltes ermöglicht wird.

Definition: Sei $P = (S, 2^S, p)$ ein endlicher Wahrscheinlichkeitsraum, mit dem Alphabet $S = \{s_1, s_2, \dots, s_m\}$ und den zugehörigen Wahrscheinlichkeiten $p(s_1), p(s_2), \dots, p(s_m)$.

$$H(P) = - \sum_{i=1}^m p(s_i) \log p(s_i)$$

heißt **Entropie** von P .

Bemerkung: Äquivalente Bezeichnungen für die in Bit gemessene Entropie sind:

- (mittlere) Unsicherheit über die Versuchsausgänge von P
- (mittlere) Unbestimmtheit von P
- (mittlerer) Informationsgehalt eines Versuchsausganges von P

3.1 Die wichtigsten Eigenschaften

Es folgt eine knappe Liste der wichtigsten Eigenschaften der Entropie [EV73]:

1. $H(P)$ ist aufgrund des Minuszeichens in der Formel stets nicht negativ.
2. Wenn ein Zeichenvorrat nur ein einziges Zeichen mit der Wahrscheinlichkeit 1, also das sichere Ereignis, besitzt, ist die Entropie $H(P) = 0$. Dadurch ist nichts ungewiss, denn a priori ist das a posteriori Ergebnis vorhersagbar.
3. Die Entropie ist maximal, wenn das Auftreten aller Symbole des Alphabets gleich wahrscheinlich ist. Diese gleichen Wahrscheinlichkeiten implizieren die größte Ungewissheit.

Folglich ist die Entropie umso geringer, je genauer unsere Kenntnis über ein System ist. In Abbildung 1 kann die Wahrscheinlichkeit von zwei unabhängigen Ereignissen s_1, s_2 , wobei $p(s_1) = 1 - p(s_2)$ ist, direkt abgelesen werden. An der Stelle $p(s_1) = p(s_2) = \frac{1}{2}$ hat die Funktion ein globales Maximum; ferner ist sie symmetrisch um diese Stelle. Vom Maximum an fällt die Funktion für $i \in \{1, 2\}$ sowohl bei $p(s_i) = 0$ als auch dem sicheren Ereignis von $p(s_i) = 1$ sehr steil zu einem Entropie-Wert von 0 ab. Dieser Zusammenhang gilt jeweils für ein Zufallsereignis. Bei mehreren Zufallsereignissen müssen die einzelnen Entropien addiert werden. Dadurch kann sie einen Entropiewert erreichen, der größer als eins ist. Die Wahrscheinlichkeit dagegen bleibt auch bei Wiederholungen definitionsgemäß immer zwischen 0 und 1.

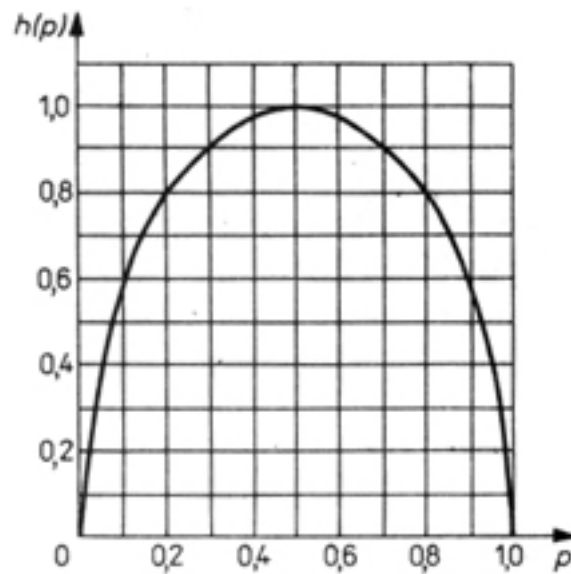


Abbildung 1: Binäre Entropiefunktion ([JJ84], Seite 62)

3.2 Maximale Entropie

Wie im Kapitel 3.1. ersichtlich ist, ist die Entropie zweier Ereignisse maximal, wenn sie gleich wahrscheinlich sind. Dies gilt auch für mehrere Ereignisse.

Satz: Sei $P = (S, 2^S, p)$ ein endlicher Wahrscheinlichkeitsraum mit dem Alphabet und Wahrscheinlichkeiten wie zuvor. Seien ferner die Ereignisse gleich wahrscheinlich, also $p(s_1) = p(s_2) = \dots = p(s_m)$, so gilt

$$H_{\max}(P) = - \sum_{i=1}^m \frac{1}{m} \log \frac{1}{m} = \log(m)$$

und die Entropie wird als **maximale Entropie** $H_{\max}(P)$ bezeichnet.

Mittels der Anzahl der Symbole des Alphabets kann die maximale Entropie, also die benötigte Anzahl an Fragen bestimmt werden, die bei gleich wahrscheinlichen Ereignissen nötig sind.

Auf den Seiten 12–15 des Buches [FK71] ist eine sehr anschauliche Beschreibung anhand eines Beispiels vorzufinden:

Soll auf einem Schachbrett mit 64 Feldern ein Feld erraten werden, ist es sinnvoll die Fragen derart zu gestalten, dass systematisch vorgegangen wird. Das zufällige Fragen nach einzelnen Feldern bringt nicht schnell das gesuchte Feld, da im ungünstigsten Fall 63 Rateversuche nötig sind, um zu erfahren, um welches Feld es sich handelt. Die sinnvollste Vorgehensweise ist die Unterteilung in jeweils zwei gleich große Bereiche. Dadurch wird die Unsicherheit bei jeder Frage um 50% verringert. Bei einem Schachbrett wären $\log(64) = 6$ Fragen nötig. „Durch eine einzige Ja-Nein-Antwort konnten ja zunächst 32, dann 16, dann 8, dann 4 und schließlich 2 und zum Schluss noch eine Wahlmöglichkeit ausgeschlossen werden“ [FK71]. Bei der Vordoppelung der Spielfelder müsste nur eine einzige weitere Frage gestellt werden, wohingegen beim unsystematischen Raten im Durchschnitt deutlich mehr Fragen gestellt werden müssten.

3.3 Relative Entropie

Die relative Unbestimmtheit einer Informationsquelle gibt den Grad der Unbestimmtheit des realisierten Zustands, bezogen auf den potenziell maximalen Informationsgehalt dieser Quelle, wider.

Definition: Sei $P = (S, 2^S, p)$ ein endlicher Wahrscheinlichkeitsraum mit dem Alphabet und Wahrscheinlichkeiten wie zuvor, dann ist

$$H_{\text{rel}}(P) = \frac{H}{H_{\text{max}}}$$

die **relative Entropie** von P .

Sie beschreibt das Verhältnis der tatsächlichen Entropie zur maximalen Entropie. Wenn die relative Entropie einer bestimmten Nachrichtenquelle zum Beispiel 0,7 ist, dann bedeutet das, dass die Quelle in der Wahl ihrer Zeichen um eine Nachricht zu bilden, etwa zu 70% so frei ist, wie es mit diesen bestimmten Zeichen überhaupt geht.

3.4 Entropie in Sprachen

Shannon „ließ eine größere Anzahl von Texten von seinen Versuchspersonen buchstabenweise durchgehen, um die Entropie der einzelnen Sprachen zu bestimmen. Dabei sollte die Versuchsperson zunächst den ersten, dann den zweiten bis zum letzten Buchstaben (einschließlich Satzzeichen und Zwischenräume) eines Textes erraten“ [FK71]. Anhand dieser Daten erstellte er den Graphen (Abbildung 2, Seite 8), der die Abhängigkeit der Entropie eines Buchstabens von der Anzahl der vorangegangenen Buchstaben als einigermaßen fehlerfreien Näherungswert darstellt. Es ist ersichtlich, dass der erste Buchstabe eine Entropie von 4,5 Bit, der fünfte 2 Bit und der sechzehnte von 1,6 Bit aufweist. Für den 100. Buchstaben eines Textes erhielt Shannon als Entropie von 1 Bit. Das bedeutet, dass der 101. Buchstabe eines Textes, bei dem 100 Buchstaben bereits gelesen wurden, durch nur einen einzigen Rateversuch entschieden werden kann.

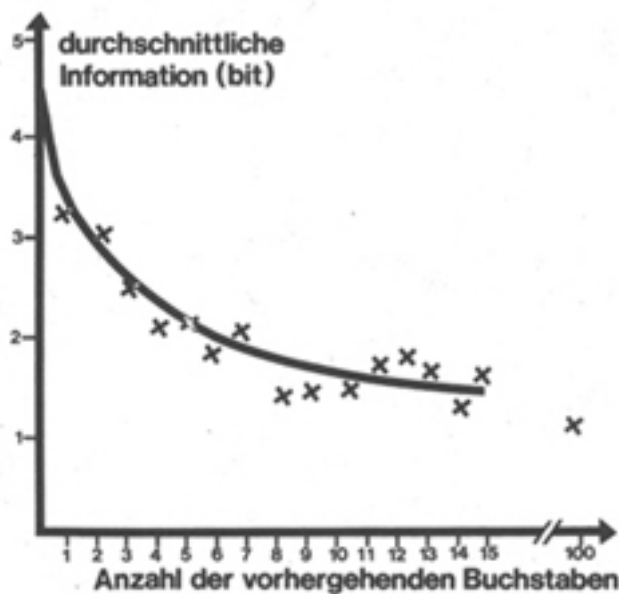


Abbildung 2: Graph nach Claude Shannon ([FK71], Seite 31)

Die Entropie für alle Sprachen, die sich aus lateinischen Buchstaben zusammensetzen, ist die Entropie $H_0 = \log 27 \approx 4,76$ Bit. Dies bedeutet, dass im Schnitt 4,76 Fragen gestellt werden müssen, um den folgenden Buchstaben zu erraten. Da bei jeder Sprache die Buchstaben mit unterschiedlichen Wahrscheinlichkeiten auftreten, muss die Entropie gesondert berechnet werden.

A.M. und I.M. Jaglom [JJ84] kamen 1967 zu dem in Tabelle 1, Seite 9 dargestellten Ergebnis.

Sprache	H (in Bit)
Englisch	4,03
Deutsch	4,10
Französisch	3,96
Spanisch	3,98

Tabelle 1: Relative Häufigkeiten der Buchstaben für die unterschiedlichen Sprachen

Es ist ersichtlich, dass die Entropie im Deutschen mit 4,10 Bit am Höchsten ist und somit die meisten Fragen gestellt werden müssen, um den nachfolgenden Buchstaben eines Textes zu erraten. Im Französischen und Spanischen hingegen sind im Schnitt weniger als 4 Fragen von Nöten.

Die Sortierung nach Wahrscheinlichkeiten des Auftretens der Buchstaben ergab nach [JJ84] die in Tabelle 2, Seite 9 dargestellten Buchstabenfolgen, wobei „-“ als Symbol für einen Zwischenraum verwendet wird.

Sprache	Buchstabenfolge
Englisch	-ETAONRI
Deutsch	-ENISTRAD
Französisch	-ESIANTUR

Tabelle 2: Buchstaben nach abnehmenden Auftritts-Wahrscheinlichkeiten geordnet

In allen betrachteten Sprachen ist der Zwischenraum das am Häufigsten auftretende Zeichen, gefolgt von dem Buchstaben „E“. Der dritt häufigste Buchstabe der drei Sprachen ist überall anders. Dies läge unter anderem daran, dass das „TH“ eine sehr typische englische Buchstabenfolge sei. Im Deutschen hingegen trete das „SCH“ verhältnismäßig oft auf. Ferner folge im Deutschen auf ein C fast immer ein H, was im Englischen und Französischen nicht der Fall sei.

4 Redundanz

4.1 Redundanz und ihre Eigenschaften

Der Begriff Redundanz kommt aus dem Lateinischen *redundare* (im Überfluss vorhanden sein). Er wurde unter anderem von Peter Fey formalisiert. Die Redundanz gibt den überflüssigen Anteil einer Nachricht an, der keine Information erhält. Sie lässt sich errechnen, indem die relative Entropie von 1 abgezogen wird.

Definition: Sei $P = (S, 2^S, p)$ ein endlicher Wahrscheinlichkeitsraum mit dem Alphabet und Wahrscheinlichkeiten wie zuvor, dann ist

$$R(P) = 1 - H_{\text{rel}}(P)$$

die **Redundanz** von P .

Mittels der Redundanz lässt sich der Verlust an Unbestimmtheit messen. Die überflüssigen Teile der Nachricht könnte man weglassen und die Nachricht wäre im Wesentlichen immer noch vollständig, beziehungsweise könnte vervollständigt werden. Die Redundanz bildet die Grundlage für die Datenkompression.

Sie ist gleich Null, wenn die Zeichen gleich wahrscheinlich und unabhängig voneinander sind. Ist dies der Fall, so „erhöht die Kenntnis eines beliebig langen Textstückes nicht die Wahrscheinlichkeit, den folgenden Buchstaben richtig zu erraten“ [JJ84].

Redundanz ist normalerweise überflüssig. Es gilt: Je redundanter ein Text ist, desto weniger Information enthält er nach Shannon. Bezogen auf die Schriftsprache sollte meiner Meinung nach zwischen der Redundanz von Buchstabenfolgen und der inhaltlichen Redundanz unterschieden werden. Die Redundanz von Buchstabenfolgen wird von hier ab in Anlehnung an die Begriffe strukturelle Information fortan als strukturelle Redundanz und die inhaltliche Redundanz als funktionale Redundanz bezeichnet.

4.2 Redundanz in Sprachen

Unter Redundanz wird, wie bereits erwähnt, die mehrfache Nennung von Informationen verstanden, die für das Verständnis des Gesamtkontexts nicht notwendig sind. Eine Sprache ist sowohl als syntaktisch als auch semantisch redundant zu bezeichnen, da sowohl Buchstaben, Silben oder Wörter als auch inhaltliche Redundanz zu verzeichnen ist.

In jeder Sprache ist diese Verteilung der Häufigkeiten von Buchstaben, Silben und Wörtern unterschiedlich. Näheres speziell zur deutschen Sprachstatistik kann in der Quelle [Mei78] nachgelesen werden. Wahrscheinliche Buchstabenabfolgen sind demjenigen, der eine Sprache gut beherrscht, verinnerlicht. Beim Korrekturlesen können Druckfehler leicht übersehen werden, da nicht die einzelnen Buchstaben, sondern die Buchstabenfolge betrachtet wird. Es ist fraglich inwieweit Schüler, die gerade das Lesen erlernt haben, oder Nicht-Muttersprachler Wörter und Sätze, in denen Fehler oder Buchstabenpermutationen auftreten, problemlos lesen können. In diesem Zusammenhang ist Redundanz hilfreich, obwohl sie als etwas Überflüssiges betrachtet wird. Wenn hingegen ein Übertragungsfehler auftritt, kann Redundanz von Vorteil sein.

4.2.1 Strukturelle Redundanz

Anhand von struktureller Redundanz ist es möglich, aus einem verstümmelten Telegramm den zu übermittelnden Text zu rekonstruieren. In den meisten Büchern wird dies anhand des folgenden Beispiels dargestellt. „Heute 17 Uhr“ wurde telegraphiert und „Heute 1. Uhr“ übermittelt. Hierbei steht der „.“ für ein aufgrund von Störung nicht übermitteltes Zeichen. Aus der übertragenen Nachricht ist schwer schließbar, um welche Uhrzeit es sich handeln sollte. Würde hingegen „Heute si..z..n Uhr“ übermittelt werden, ist die Nachricht noch entzifferbar. Die Semantik der Nachricht wurde nicht geändert. Lediglich längere Worte wurden verwendet. Daraus ergibt sich, dass aufgrund dieses Überflusses Mitteilungen trotz der vorhandenen Lücken verständlich bleiben. Die Kenntnis des Zusammenhanges in Bezug auf die Uhrzeit ist in diesem Beispiel hilfreich. Es lässt sich in die heutige Zeit übertragen, in dem ein übermitteltes Fax betrachtet wird, bei dem die Tonerfarbe zum Teil schwer bis gar nicht lesbar ist. Da der inhaltliche Zusammenhang für diesen Vorgang zur Dechiffrierung von Nöten ist, ist die strukturelle Redundanz nicht losgelöst von der funktionalen Redundanz zu betrachten.

Ein ähnliches Beispiel stellen Musiker dar, die beim Vomblattspielen durch flüchtigen Blick einen Komplex von mehreren Noten erkennen können, das. Sie müssen nicht Note für Note lesen. Ein grober Blick genügt meist, vor allem wenn ein Stück eines Künstlers gespielt werden soll, von dem sie bereits mehrere Stücke spielten. Sowohl epochentypisch wiederkehrende Tonfolgen als auch der Stil des Komponisten sind dem Musiker hilfreich.

4.2.2 Funktionale Redundanz

Die Vorgehensweise des Einarbeitens in einen nicht vertrauten wissenschaftlichen Zusammenhang wird meist von Redundanz geprägt. Hierbei handelt es sich jedoch um Redundanz in semantischer Hinsicht. Häufig werden unterschiedliche Bücher zu Rate gezogen, die in verschiedenster Weise die Thematik erklären. Dadurch ist ein Einstieg in ein fremdes Gebiet leichter, da Inhalte öfters repetiert und mit anderen Worten ausgedrückt werden. „Denn hier haben wir es mit so komplizierten Zusammenhängen zu tun, dass wir nur durch die kritische Konfrontation verschiedener Darstellungen zu einem Gesamtbild des Sachverhaltes gelangen können.“ [Sei68]

Zunächst wird die redundante strukturelle Information gelesen und dann interpretiert. Daraus entsteht funktionale Information beziehungsweise durch das Repetieren der Thematik funktionale Redundanz. Sollte diese Redundanz dafür sorgen, dass mehr verstanden wurde, müsste es eine pragmatische Redundanz geben, die die Wirkung auf die redundante Information darstellt.

5 Veranschaulichung der pragmatischen Information nach von Weizsäcker

Die pragmatische Information ist nach von Weizsäcker abhängig von Erstmaligkeit und Bestätigung. Sie ist minimal, wenn auch die Entropie minimal ist. Dies ist unter anderem der Fall, sobald das Ereignis sicher ist beziehungsweise „die Information bereits vollständig bekannt, also redundant ist“ [EFS98]. Da das Ereignis sicher ist, ist keine Unbestimmtheit vorhanden und es kann vorhergesagt werden, was eintreten wird. Folglich betragen die Bestätigung 100% und die Erstmaligkeit 0%. Daraus lässt sich schließen, dass die pragmatische Information minimal ist, wenn die Information bekannt, also redundant ist. Dieser Sachverhalt ist in der Abbildung 3, Seite 13 am Koordinatenursprung wieder zu finden.

Ebenfalls minimal ist die pragmatische Information, wenn die Information neu und unverständlich ist. Dies entspricht einer Erstmaligkeit von 100% und einer Bestätigung von 0%. Die Information knüpft nicht an bereits bekanntes Wissen an. Da sie vollkommen neu und damit unverständlich ist, wird auch keine Wirkung beim Empfänger erzielt.

Im Gegensatz dazu ist die Shannonsche Informationsentropie hingegen ist dann maximal, wenn die Erstmaligkeit 100% und minimal wenn die Erstmaligkeit 0% beträgt.

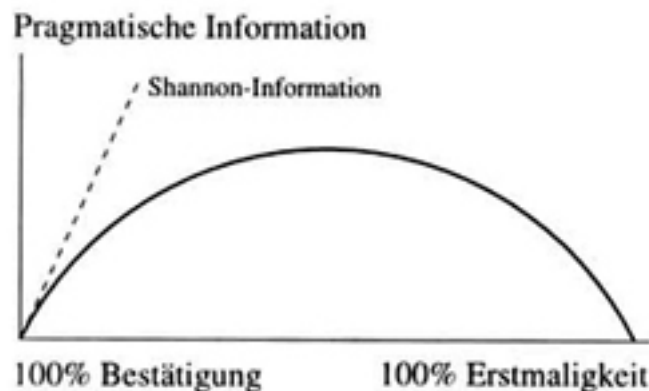


Abbildung 3: Schematische Darstellung der pragmatischen Information ([EFS98], Seite 60)

Bei 50% Bestätigung und 50% Erstmaligkeit ist die pragmatische Information maximal. „E. v. & C. v. Weizsäcker (1972) haben die These vertreten, daß lebende Systeme zwischen beiden Grenzfällen und damit in der Nähe des Maximums der pragmatischen Information operieren. Das heißt, die für die Evolution bedeutsame Information muß einerseits einen gewissen Neuheitswert haben, der andererseits aber auf der Grundlage der bereits vorhandenen Information auch verstanden werden kann.“ [EFS98]

6 Zusammenfassung

Wie eingangs erwähnt, wurde die Entscheidung getroffen, einen Überblick über den Themenkomplex zu geben. Daher sind die grundlegenden Aspekte der Informationstheorie behandelt und anhand von Beispielen erläutert worden.

Die wichtigsten Ergebnisse der Arbeit sind die Folgenden:

In Kapitel 2 wurden die Begriffe gebundene und freie beziehungsweise strukturelle, funktionale und pragmatische Information behandelt und erläutert. Der informationstheoretische Aspekt der Information ist anhand eines Beispiels von dem intuitiven Verständnis differenziert worden.

Das Kapitel 3 befasste sich mit der Entropie und ihrer Spezifikation in maximale und relative Entropie. Anhand dieser Begriffsbildung wurde es möglich, den Informationsgehalt der Sprache zu betrachten. Es ist meiner Meinung nach sehr interessant, den Vergleich der unterschiedlichen Sprachentropien wahrzunehmen und zu erfahren, dass die deutsche Sprache von den getesteten Sprachen die höchste Entropie aufweist. Daraus könnte gefolgert werden, dass in der deutschen Sprache sehr viel redundante Buchstaben verwendet werden.

Das 4. Kapitel thematisiert die Redundanz. Hier habe ich versucht, den Begriff der strukturellen und funktionalen Information auf die Redundanz zu übertragen. Diese Begriffe wurden mittels Beispielen erläutert.

Das Kapitel 5 hat den Begriff pragmatische Information an einem Beispiel von C. F. von Weizsäcker dargestellt und dadurch den Zusammenhang zwischen pragmatischer Information und dem Shannonschen Informationsbegriff erläutert.

Bibliographie

- [EFS98] EBELING, W. ; FREUND, J. ; SCHWEITZER, F.: *Komplexe Strukturen: Entropie und Information*. Stuttgart : B. G. Teubner, 1998
- [EV73] ENGEL, Ulrich ; VOGEL, Irmgard: Statistik für Linguisten – Eine methodische Darstellung / Institut für Deutsche Sprache. Mannheim : Institut für Deutsche Sprache, 1973. – Forschungsbericht
- [FK71] FRÖHLICH, Werner D. ; KOSZYK, Kurt: *Die Macht der Signale - Information, Kommunikation und Gesellschaft*. Hamburg : Rowohlt, 1971
- [JJ84] JAGLOM, A. M. ; JAGLOM, I. M.: *Wahrscheinlichkeit und Information*. VEB - Deutscher Verlag der Wissenschaften, 1984
- [Mei78] MEIER, Helmut: *Deutsche Sprachstatistik*. Hildesheim : Georg Olms Verlag, 1978
- [Sei68] SEIFFERT, Helmut: *Information über die Information*. München : Verlag C. H. Beck, 1968