

2 Information: Ereignis, Wahrscheinlichkeit, Entropie und Sinngehalt

In Kapitel 1 wurde im Abschn. 1.1 ausgeführt, dass Kommunikation in der Übergabe oder im Austausch von Nachrichten besteht und dass der Empfang von Nachrichten auf Wahrnehmung von Signalen beruht. Signale sind nur die Träger von Nachrichten. Um an die eigentliche Nachricht (d.h. an die Bedeutung oder den Sinngehalt) zu kommen, muss das Signal vom Empfänger des Signals interpretiert werden. „Interpretation“ heißt „dem Signal eine Bedeutung (Sinngehalt) zuordnen“. Der vom Empfänger zugeordnete Sinngehalt sollte mit dem vom Sender gemeinten Sinngehalt möglichst genau übereinstimmen. Wahrnehmung und Interpretation beschränken sich aber nicht nur auf solche Signale, die zum Zweck der Kommunikation erzeugt werden, sondern erstrecken sich auch auf andere Signale, die von sonstigen Ereignissen und Erscheinungen herrühren und ebenfalls oft Bedeutungen tragen. Dieser letztgenannte Punkt wird in Abschn. 2.2.2 kurz angesprochen.

Unterschiede zwischen dem vom Empfänger zugeordneten Sinngehalt und dem vom Sender gemeinten Sinngehalt können dadurch entstehen, dass das vom Sender gesendete Signal beim Empfänger beschädigt ankommt und deshalb vom Empfänger falsch interpretiert wird. Es kann aber auch sein, dass selbst ein Signal, das unverändert beim Empfänger ankommt, vom Empfänger falsch interpretiert wird, weil es sich beim Sender und Empfänger um verschiedene Individuen oder Einrichtungen handelt, die nicht identisch sind.

In diesem 2. Kapitel interessiert nur der Fall der *einseitig* gerichteten sinnngenauen Kommunikation zweier Kommunikationspartner KP_1 und KP_2 , bei der KP_1 nur sendet und KP_2 nur empfängt, und bei der das *Signal ungestört* und unverändert übertragen und von KP_2 richtig interpretiert wird. Zentraler Gegenstand dieses Kapitels ist die Bestimmung der *Höhe an Information*, die der Empfänger KP_2 dabei erhält. [Auf die Einflüsse von Störungen und Verzerrungen des übertragenen Signals wird im 4. Kapitel eingegangen. Die bei beidseitiger Nachrichtenübertragung zwischen zwei Partnern durch Rückkopplungen möglichen Effekte, die in Abschn. 1.6 behandelt wurden und auf die in späteren Kapiteln erneut eingegangen wird, werden in diesem 2. Kapitel ebenfalls nicht betrachtet].

2.1 Erste Erläuterungen und Übersicht

Mit seiner bahnbrechenden Publikation „A Mathematical Theory of Communication“ hat C. E. Shannon (1948) die moderne Informationstheorie begründet, die heute in der elektrischen Nachrichtentechnik beim Entwurf moderner Codier- und Übertragungssysteme eine maßgebende Rolle spielt und sich dort sehr bewährt. Eines von mehreren grundlegenden Ergebnissen, die Shannon in seiner Arbeit hergeleitet hat, ist eine Formel für den wahrscheinlichkeitstheoretischen Erwartungswert der Informationshöhe, die eine Informationsquelle liefert. Weil diese Formel die gleiche Gestalt hat wie die Formel für die Entropie in der Quantenphysik, hat Shannon den Erwartungswert der Informationshöhe

ebenfalls als Entropie bezeichnet. Die Informationshöhe gibt nur an, wie viel Information geliefert wird. Sie sagt nicht aus, wie wichtig die Information ist und auch nicht, welche Aussagen und welchen Sinngehalt die Information beinhaltet.

Da die quantenphysikalische Entropie aus der von Clausius eingeführten thermodynamischen Entropie von Gasen abgeleitet ist, und die thermodynamische Entropie wiederum angibt, in welchem Maß Wärme in mechanische Energie umgewandelt werden kann, kam es zu einer Fülle von Diskussionen darüber, was Information überhaupt ist, ob es eine geistig-immaterielle Größe oder eine energetisch-materielle Größe ist. Genannt seien hier (a) der Physiker L. Brillouin (1950), der anhand des sogenannten Maxwell'schen Dämons den Zusammenhang zwischen der informationstheoretischen Entropie von Shannon mit der Entropie in der Physik untersucht hat, (b) der Philosoph P. Janich (2006), der in seinem Buch „Was ist Information?“ die Darstellung von Shannon wie auch übliche Interpretationen der Physik wegen mangelhafter Unterscheidungen scharf kritisiert, für teilweise unzutreffend hält und Vorschläge zur Reparatur der Informationstheorie macht (die aus Sicht der Nachrichtentechnik aber nicht praktikabel sind) und (c) der Informatiker W. Gitt (2002), der in seinem Buch „Am Anfang war die Information“ darlegt, dass das, was Shannon beschreibt und quantifiziert, nicht immer Information sei, und selber Bedingungen für das Vorhandensein von Information formuliert.

Der Informationsbegriff ist also äußerst vielschichtig. Nicht von ungefähr schreibt J. Reischer (2006): „Wer über Information redet oder schreibt, sticht in ein Hornisssennest und setzt sich anschließend darauf.“ Der Verfasser der vorliegenden Abhandlung denkt, dass ihm das nicht passiert. Er ist bemüht, verschiedene Standpunkte, die ja nicht in jeder Hinsicht abwegig sind, miteinander zu versöhnen und den Informationsbegriff zu entmystifizieren. Als Nachrichtentechniker ist er allerdings der Meinung, dass die Informationstheorie von Shannon, die zugleich die Grundlage für ein nützliches Einordnungsschema liefert, ein unverzichtbarer Bestandteil einer allgemeinen Theorie der Kommunikation sein muss.

Ein wesentliches Konzept der Shannon'schen Informationstheorie basiert auf einer Auffassung, die auch in der natürlichen Umgangssprache zum Ausdruck kommt: Wenn jemandem eine Begebenheit mitgeteilt wird, die er bereits kennt, dann sagt er, dass diese Mitteilung ihm keine Information bringe. Information muss also etwas *Neues* liefern. Etwas *Neues* allein reicht aber noch nicht aus. Man muss das Neue auch irgendwie einordnen können und verstehen. Das Neue muss also auch im eigenen Kopf latent (d.h. verborgen) vorhanden sein und durch die Mitteilung ins Bewusstsein geholt werden. Eine Mitteilung in einer völlig fremden Sprache, von der man auch nicht das Geringste versteht, bringt einem keine Information (wohl aber kann sie einem anderen, der die Sprache versteht, Information bringen). Auch dieser Aspekt, das etwas, das Information liefert, auf der Empfangsseite bereits latent vorhanden sein muss, bildet eine Basis der Shannon'schen Informationstheorie. Vor dem Empfang einer Mitteilung herrscht auf der Empfangsseite eine Unsicherheit über das, was kommen wird. Mit dem Eintreffen der Mitteilung wird diese Unsicherheit beseitigt. Das Maß an beseitigter Unsicherheit ist nach Shannon ein Maß für die Höhe der empfangenen Information (vergl. hierzu Abschn. 1.3.3 über Prädiktor-Korrektor-Verfahren).

Die vorstehend genannten Zusammenhänge gelten nicht nur für den menschlichen Empfänger von Information sondern auch für den maschinellen Empfänger von Information. Damit das Neue eingeordnet werden kann, muss es auch in der Maschine bereits latent vorhanden sein. Beim Empfang der Mitteilung wird das Neue zwar nicht ins Bewusstsein geholt, das

eine Maschine^{2.1} nach Ansicht des Verfassers nicht besitzt, sondern von der Maschine für die Durchführung einer selektiven Aktion benutzt. Diese selektive Aktion kann auch in der Ausgabe einer modifizierten Mitteilung an eine höhere Instanz bestehen. In der höheren Instanz ist eventuell ein Bewusstsein vorhanden.

Ein zusätzlicher Aspekt ist folgender: Wenn dieselbe Mitteilung zwei empfangende Kommunikationspartner KP_{2a} und KP_{2b} erreicht, von denen der eine den Inhalt bereits kennt, aber der andere noch nicht, dann erhält der eine Kommunikationspartner keine Information, während der andere dagegen sehr wohl Information erhält. Dies macht deutlich, dass Information vom individuellen Vorwissen abhängt und damit eine subjektive Größe ist. Diesen Aspekt hat Shannon in seiner oben genannten Publikation zwar nicht behandelt, er lässt sich aber unschwer in der Shannon'schen Informationstheorie berücksichtigen.

Nur in solchen Fällen, bei denen jeder empfangende Kommunikationspartner das gleiche Vorwissen besitzt, lässt sich Information als etwas Objektives darstellen, das für jeden empfangenden Kommunikationspartner gleich ist. Shannon hat seine Formel für die Entropie (von Information) am Beispiel der maschinellen Übertragung von Buchstaben hergeleitet. Gleiche maschinelle Empfänger besitzen bezüglich des Empfangs von Signalen, die Buchstaben bedeuten, das gleiche „Vorwissen“. Die an diesem Beispiel entwickelte Formel für die Entropie von Information bezieht sich auf die Interpretationsebene der Maschine. Es handelt sich also um die *von der Maschine erfahrene* Entropie von Information.

Die von Shannon entwickelte Formel für die Entropie bestimmt nur ein quantitatives Maß. Aus der Formel geht nicht hervor, dass es sich um das Auftreten von Buchstaben handelt. Die Formel gilt allgemein für das Auftreten beliebiger Ereignisse. Was bei Shannon als „Information“ bezeichnet wird, ist genau genommen die bereits genannte „Informationshöhe“. Anhand welcher speziellen Art von Ereignissen die Informationshöhe berechnet wird, muss der jeweiligen Gegebenheit entnommen werden. Die von Shannon betrachteten Ereignisse bestehen im Auftreten einzelner Buchstaben.

Shannon weist in seiner eingangs genannten Publikation ausdrücklich darauf hin, dass eine Buchstabenfolge für den in einer höheren Ebene angesiedelten menschlichen Benutzer der Maschine normalerweise auch eine semantische Bedeutung hat, die aber für den maschinellen Empfang von Buchstaben in der tiefer gelegenen Ebene ohne Belang ist. Eine auf einer höheren Ebene durchgeführte Interpretation von solchen Signalen, die von einer tiefer gelegenen Ebene hochgereicht wurden, hat es im Normalfall mit einer Entropie zu tun, die verschieden ist von derjenigen in der darunter gelegenen Ebene. Festzuhalten ist daher

[2.1] Entropie, Interpretationsebene und Information:

Der in der Informationstheorie von Shannon benutzte Begriff „Entropie“ (das ist der wahrscheinlichkeitstheoretische Erwartungswert der von einer Informationsquelle gelieferten „Informationshöhe“) bezieht sich immer auf eine bestimmte Interpretationsebene. Die Formel für die Entropie liefert nur ein quantitatives Maß, unabhängig vom Sinngehalt, vergleichbar mit einer Waage, die nur ein Gewicht liefert unabhängig von dem, was (d.h. ob Sand oder Gold oder Wasser oder sonst etwas) gewogen wird. Zur vollständigen Angabe von Information gehört neben der Informationshöhe auch die Angabe des Sinngehalts.

^{2.1} Der Philosoph Th. Metzinger (2009) ist der Ansicht, dass auch Maschinen ein Bewusstsein haben können. Das hält der Verfasser dieser Abhandlung für abwegig. Maschinen (auch Computer) sind nur Werkzeuge des Menschen, siehe Abschn. 1.4.2 sowie 5.2 und 5.3. Sie werden nie aus originär eigener Initiative aktiv.

Ein Anliegen dieser Abhandlung besteht darin, zu zeigen, dass die dieselbe Informationstheorie sich auf beliebige Interpretationsebenen anwenden lässt.

Versuche, den Begriff „Information“ in einer solchen Weise zu definieren, dass Information etwas Objektives darstellt, das vom Vorwissen der (empfangenden) Beobachter unabhängig ist, oder dass Information nur dann vorliegt, wenn der Mensch einen Sinn im Signal entdeckt, stoßen entweder irgendwo auf Widerspruch oder erzwingen unangemessene Einschränkungen. Solche nicht angemessenen Einschränkungen machen z.B. Janich und Gitt.

So viel zu den ersten (noch groben) Erläuterungen! Nun zum Überblick über die weiteren Einzelheiten, die in diesem 2. Kapitel behandelt werden:

In den nachfolgenden Unterabschnitten dieses Abschnitts 2.1 werden an einem einfachen Beispiel für die Übertragung von Bedeutungen die Begriffe Informationshöhe, Entropie und weitere veranschaulicht und die Rolle von Interpretationsebenen kurz diskutiert. Genauere Erläuterungen zur Informationshöhe folgen in den Abschnitten 2.2 und 2.4. Dabei wird zwangsläufig von der Wahrscheinlichkeitstheorie und der Mengentheorie Gebrauch gemacht. Damit die Abschnitte 2.2 und 2.4 und die weiteren Abschnitte dieses 2. Kapitels auch von Lesern ohne Vorkenntnisse von Wahrscheinlichkeitsbeziehungen und Mengen in voller Länge gelesen werden können, werden im Abschnitt 2.3 die wichtigsten Zusammenhänge der Wahrscheinlichkeitstheorie ausführlich und in möglichst einfacher Weise eingeführt. In Abschn. 2.5 wird dann auf die Shannon'sche Formel für die Entropie, deren Herleitung, Eigenschaften und Anwendungen näher eingegangen. Nach einer Erläuterung der Komplexität von Information im Abschn. 2.6 folgen im Abschn. 2.7 eine Beschreibung der höheren Interpretationsebenen und einige allgemeinere Betrachtungen über Information. Dort wird sich zeigen, dass insbesondere die subjektive Bewertung von Information auf die eigenständigen Felder der Psychologie und Werbetechniken führen.

Die vielfältigen Bezüge zwischen informationstheoretischer Entropie und thermodynamischer Entropie und mechanischer Energie werden in einem eigenen 3. Kapitel behandelt. Dort wird auf die formal mathematischen Zusammenhänge dieser Größen detailliert eingegangen.

2.1.1 Zur Übertragung von Bedeutungen und zum Informationsbegriff

Das zentrale Ziel der eingangs genannten Publikation von Shannon war die Klärung der Frage, wie der sendende Kommunikationspartner KP_1 die Zuordnung des Signals zweckmäßiger Weise gestaltet, damit die Übertragung der Bedeutung und die Wiedergewinnung der Bedeutung durch Interpretation des Signals durch den Kommunikationspartner KP_2 auf der Empfangsseite möglichst effizient erfolgen.

Ausgangspunkt ist die Annahme (Voraussetzung), dass beide Kommunikationspartner über eine gleiche Liste von endlich vielen verschiedenen Bedeutungen S_i mit $i = 1, 2, \dots, N$ verfügen [im Unterschied zu Abb. 1.2 besitzen jetzt beide Partner gleiche Vorstellungswelten]. Der sendende Kommunikationspartner KP_1 wählt aus dieser Liste eine bestimmte Bedeutung S_k aus, die er dem empfangenden Kommunikationspartner KP_2 mitteilen will, und ordnet dieser Bedeutung S_k ein Signal s_k zu, das er an den empfangenden Kommunikationspartner KP_2 sendet. Der empfangende Kommunikationspartner KP_2 bestimmt anhand des empfangenen Signals s_k die in seiner Liste enthaltene zugehörige Bedeutung S_k .

Die empfangene Information besteht für den Kommunikationspartner KP_2 in dem *neu* erhaltenen Wissen darüber, welche Bedeutung auf der Sendeseite gemeint ist.

Ein umfangreicherer Kommunikationsprozess besteht in der Regel darin, dass der sendende Kommunikationspartner KP_1 nacheinander eine ganze Folge von Bedeutungen S_i aus seiner Liste auswählt und eine entsprechende Folge von Signalen s_i an den empfangenden Kommunikationspartner KP_2 sendet. Die Natur üblicher Kommunikationsprozesse bringt es dabei mit sich, dass gewisse Bedeutungen häufiger ausgewählt werden und andere weniger. Eine Untersuchung darüber, wie die Auswahl zweckmäßiger Weise zu gestalten ist, erweist sich als sehr aufschlussreich. Sie liefert zugleich auch Auskunft über die Höhe von Information. Es zeigt sich, dass die Übertragung von Bedeutungen dann am effizientesten ist, wenn für beliebige Folgen von Bedeutungen die Auswahl der Bedeutungen auf der Sendeseite im Mittel mit dem geringsten Aufwand erfolgt. Was das im Einzelnen bedeutet, sei nun anhand von Abb. 2.1 näher erläutert.

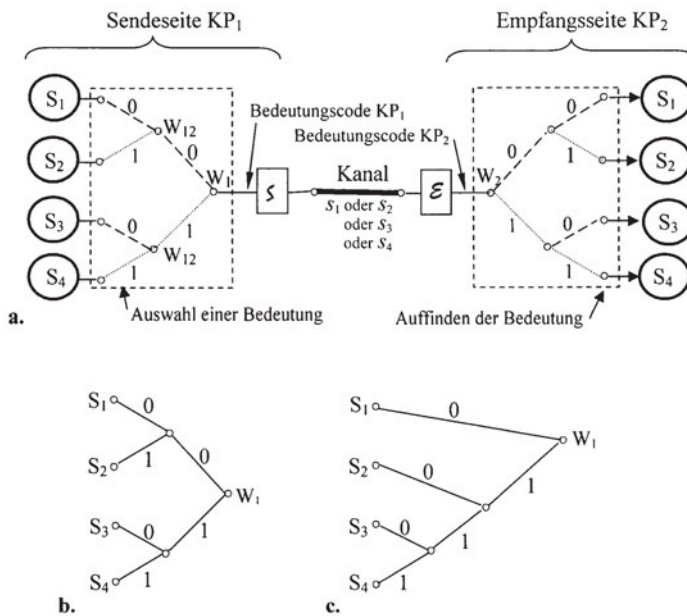


Abb. 2.1 Übertragung von nur vier möglichen Bedeutungen S_1, S_2, S_3 und S_4 über einen Kanal (oder Medium)

- Auswahl einer Bedeutung S_i für $i = 1, 2, 3, 4$ mittels zweier Binärentscheidungen und Zuordnung des zugehörigen Signals s_i
- Codebaum in Abb. a
- Alternativer Codebaum

Abb. 2.1 zeigt den primitiven Fall, dass der sendende Kommunikationspartner KP_1 und der empfangende Kommunikationspartner KP_2 dieselben nur vier verschiedenen (scharfen) Bedeutungen S_1, S_2, S_3 und S_4 kennen. Dies stellt eine erhebliche Vereinfachung im Vergleich zu den Betrachtungen in Abschn. 1.5.3 dar. Mit S_1 sei z.B. die Bedeutung *Habe Hunger*, mit S_2 die Bedeutung *Lass mich in Ruhe*, mit S_3 die Bedeutung *Möchte spielen* und mit S_4 die Bedeutung *Habe Schmerzen* gemeint. Mit S_1 bis S_4 können aber auch andere

Bedeutungen gemeint sein, z.B. vier verschiedene Buchstaben. Die nachfolgenden Betrachtungen gelten unabhängig von der Art der Bedeutungen (ob einzelner Buchstabe oder irgendein umfangreicher Sinngehalt). Die Bedeutungen sollen in willkürlicher Reihenfolge ausgewählt werden können, eine Grammatik über die Zulässigkeit von Reihenfolgen soll hier nicht vorhanden sein. Von Grammatiken und von der Art, wie Bedeutungen gespeichert werden, wird erst im 5. Kapitel und besonders im 6. Kapitel die Rede sein.

Zunächst zum sendenden Kommunikationspartner KP_1 :

Wenn der sendende Kommunikationspartner KP_1 , z.B. ein Hund, eine der oben genannten Bedeutungen akustisch ausdrückt, dann geschieht das mit einem zweiseitigen Vorgang.

Der erste Teil besteht in der *Auswahl* der zu artikulierenden Bedeutung aus der Liste der vier möglichen Bedeutungen. Dieser Auswahlvorgang erfolgt in zwei Schritten, die im gestrichelt eingerahmten Teil der linken Hälfte von Abb. 2.1a dargestellt sind. Ausgehend vom Punkt W_1 wird im ersten Schritt zwischen den Möglichkeiten 0 und 1 und anschließend vom jeweils erreichten Punkt W_{12} im zweiten Schritt wieder zwischen den Möglichkeiten 0 und 1 ausgewählt. Auf diese Weise wird z.B. die Bedeutung S_3 über die Auswahlfolge (Codewort) 10 erreicht. Die vier möglichen Codewörter bei der Auswahl stellen den Bedeutungscode KP_1 auf der Sendeseite dar:

$$00 \rightarrow S_1; 01 \rightarrow S_2; 10 \rightarrow S_3; 11 \rightarrow S_4 \quad (2.1)$$

Der zweite Teil des Vorgangs, der gleichzeitig mit dem ersten Teil ablaufen kann (aber nicht muss) besteht darin, dass jedem Codewort des Bedeutungscode KP_1 ein Signal s_i zugeordnet wird, z.B.

$$00 \rightarrow s_1; 01 \rightarrow s_2; 10 \rightarrow s_3; 11 \rightarrow s_4 \quad (2.2)$$

Diese Signalzuordnung geschieht im Artikulationsorgan $\mathcal{S}$. Die vier möglichen Signale s_1 bis s_4 in können im Prinzip beliebige Formen haben, sie müssen sich lediglich voneinander unterscheiden, damit sie von einem empfangenden Kommunikationspartner KP_2 richtig interpretiert werden können. Weiter unten wird deutlich werden, dass nicht das Signal selbst, sondern der mit dem Signal festgelegte Auswahlvorgang für die Höhe einer Information entscheidend ist.

Nun zum empfangenden Kommunikationspartner KP_2 :

In Abb. 2.1a wird vorausgesetzt, dass der empfangende Kommunikationspartner KP_2 das Signal s_i über einen idealen Kanal (Medium) unverfälscht erhält. Der Interpretationsvorgang auf der Empfangsseite setzt sich ebenfalls aus zwei Teilen zusammen.

Der erste Teil besteht darin, dass im Empfangsorgan $\mathcal{E}$ jedem empfangenen Signal s_i ein Codewort des empfangsseitigen Bedeutungscode KP_2 zugeordnet wird. Im einfachen Fall von Abb. 2.1a ist das die umgekehrte Zuordnung von (2.2)

$$s_1 \rightarrow 00; s_2 \rightarrow 01; s_3 \rightarrow 10; s_4 \rightarrow 11 \quad (2.3)$$

Der zweite Teil besteht im Auffinden der jeweils zugehörigen Bedeutung. Das geschieht mit zwei binären Entscheidungsschritten, die im gestrichelt eingerahmten Teil der rechten Hälfte von Abb. 2.1a dargestellt sind. Die vier verschiedenen Codewörter oder Folgen von 0 und 1 bestimmen jeweils den Pfad zur zugehörigen Bedeutung.

A: Einheit „bit“ als Maß für die Anzahl binärer Auswahlsschritte und Entscheidungen

Für die weitere Diskussion ist die Einführung des Begriffs „bit“ zweckmäßig. Sowohl ein binärer Auswahlsschritt auf der Sendeseite als auch eine binäre Entscheidung auf der Empfangsseite werden kurz als bit (von binary digit) bezeichnet. Ein bit kann nur den Wert 0 oder 1 haben:

$$1 \text{ bit} = \begin{cases} 0 \\ \text{oder} \\ 1 \end{cases} \quad (2.4)$$

Die in den Beziehungen (2.1) bis (2.3) aufgetretenen Codewörter bestehen alle aus je 2 bit. Das (klein geschriebene Wort) bit ist eine abstrakte (gedankliche) Größe, die durch ein (groß geschriebenes Wort) Bit physikalisch dargestellt und übertragen werden kann, siehe dazu Abb. 1.3 im 1. Kapitel.

B: Unterschiedliche Auswahl von Bedeutungen über verschiedene Codebäume

Statt den Bedeutungen S_1 bis S_4 direkt entsprechende Signale s_1 bis s_4 umkehrbar zuzuordnen, wurden in Abb. 2.1a mit den gestrichelt eingerahmten Teilen die BedeutungsCodes KP_1 und KP_2 zwischengeschaltet, weil dies eine einfache Beschreibung des Auswahlvorgangs erlaubt. Darüber hinaus lassen sich anhand der BedeutungsCodes Aussagen über die Effizienz der Übertragung von Bedeutungen und auch zu deren Informationshöhe machen.

Die gestrichelt eingerahmten Teile in Abb. 2.1a werden *Codebäume* genannt. Der auf der Empfangsseite gezeichnete Codebaum ist die gespiegelte Version des Codebaums auf der Sendeseite. In Abb. 2.1b ist dieser Codebaum nochmal separat dargestellt. Für die Bildung von Codebäumen gibt es mehrere Möglichkeiten. Einen alternativen Codebaum für die Auswahl einer Bedeutung (bzw. deren Auffindung) zeigt Abb. 2.1c. Bei diesem erfolgt die Auswahl der Bedeutung S_1 in einem Schritt, die Auswahl der Bedeutung S_2 in zwei Schritten und die Auswahl der Bedeutungen S_3 und S_4 in jeweils drei Schritten. Statt des BedeutungsCodes (2.1) gilt jetzt

$$0 \rightarrow S_1; \quad 10 \rightarrow S_2; \quad 110 \rightarrow S_3; \quad 111 \rightarrow S_4 \quad (2.5)$$

1. Zwischenbemerkung:

Obwohl die einzelnen Codewörter aus unterschiedlich vielen bit bestehen, lässt sich aus jeder beliebigen Folge von bit eindeutig auf die Folge der Codewörter und der zugehörigen Bedeutungen schließen. Das liegt daran, dass es in (2.5) nur ein einziges Codewort gibt, das mit 0 beginnt und alle übrigen Codewörter mit 1 beginnen, und dass ein mit 1 beginnendes Codewort entweder mit einer 0 endet oder nach drei aufeinander folgenden 1.

$$\text{Beispiel: } 011110110010 \rightarrow S_1 S_4 S_2 S_3 S_1 S_2 \quad (2.6)$$

C: Zur Effizienz der Übertragung von Bedeutungen

Zur Beurteilung von Effizienz beim Senden wird jetzt der Fall betrachtet, dass bei mehreren aufeinanderfolgenden Kommunikationsprozessen oder bei einem einzelnen umfangreicheren Kommunikationsprozess unter z.B. 80 ausgewählten Bedeutungen 40 mal die Bedeutung S_1 und 20 mal die Bedeutung S_2 und je 10 mal die Bedeutungen S_3 und S_4 auftreten.

Legt man für die Auswahl den Codebaum in Abb. 2.1b zugrunde, dann ergibt das beim zugehörigen Bedeutungscode KP_1 eine Folge von 160 bit, weil jede Bedeutung mit zwei bit codiert wird. Legt man dagegen für die Auswahl den Codebaum in Abb. 2.1c zugrunde, dann erhält man beim zugehörigen Bedeutungscode KP_1 eine Folge von nur 140 bit, weil die 40 mal auftretende Bedeutung S_1 mit je nur einem bit, die 20 mal auftretende Bedeutung S_2 mit je zwei bit und die je nur 10 mal auftretenden Bedeutungen S_3 und S_4 mit je drei bit codiert werden. Der Codebaum in Abb. 2.1c ist für diesen Fall effizienter, weil damit die gleiche Anzahl von Bedeutungen mit weniger bit codiert und übertragen wird. Die im Mittel aufgewendete Anzahl von bit beträgt 2 bit pro Bedeutung beim Codebaum in Abb. 2.1b und nur $1\frac{3}{4}$ bit/Bedeutung beim Codebaum in Abb. 2.1c. Hierbei ist wichtig zu vermerken, dass die Werte 2 bit pro Bedeutung und $1\frac{3}{4}$ bit pro Bedeutung für die im vorigen Absatz genannten unterschiedlichen Häufigkeiten der einzelnen Bedeutungen gelten.

Anders ist das nämlich im Fall, dass bei einem umfangreicheren Kommunikationsprozess in unter z.B. 80 ausgewählten Bedeutungen die vier Bedeutungen S_1 bis S_4 gleich häufig, also je 20 mal, auftreten. Bei Zugrundelegung des Codebaums in Abb. 2.1b ergibt das beim zugehörigen Bedeutungscode KP_1 wieder eine Folge von 160 bit, beim Codebaum in Abb. 2.1c hingegen eine Folge von 180 bit, weil 20 mal ein bit, 20 mal zwei bit und 40 mal drei bit gebraucht werden. In diesem Fall, bei dem alle Bedeutungen mit gleicher Häufigkeit ausgewählt werden, ist also der Codebaum in Abb. 2.1b effizienter. Die pro Bedeutung im Mittel aufgewendete Anzahl von bit beträgt, wie vorher, 2 bit/Bedeutung beim Codebaum in Abb. 2.1b, jedoch $2\frac{1}{4}$ bit/Bedeutung beim Codebaum in Abb. 2.1c.

Wie später in Abschn. 2.5.4 noch näher ausgeführt wird, ist der Codebaum in Abb. 2.1b optimal, wenn die Auswahl aller vier Bedeutungen mit gleicher Wahrscheinlichkeit P erfolgt, wenn also $P(S_1) = P(S_2) = P(S_3) = P(S_4) = \frac{1}{4}$. Im Unterschied dazu ist der Codebaum in Abb. 2.1c optimal, wenn die Auswahl der Bedeutungen mit den Wahrscheinlichkeiten $P(S_1) = \frac{1}{2}$, $P(S_2) = \frac{1}{4}$, $P(S_3) = P(S_4) = \frac{1}{8}$ erfolgen. Optimal bedeutet, dass es keinen Codebaum gibt, bei dem die mittlere Anzahl der bit/Bedeutung geringer ist als bei den Codebäumen in Abb. 2.1b und Abb. 2.1c.

2. Zwischenbemerkung:

Der Begriff *Wahrscheinlichkeit* hat mit Zufälligkeit zu tun. Zufällig ist die Auswahl nur aus der Sicht des empfangenden Kommunikationspartners KP_2 . Der sendende Kommunikationspartner KP_1 trifft die Auswahl in der Regel nicht zufällig, sondern aufgrund einer Absicht, die aber der empfangende Kommunikationspartners KP_2 zunächst nicht kennt. Die aus der beobachteten Häufigkeit abgeleitete Wahrscheinlichkeit wird als *a-posteriori*-Wahrscheinlichkeit bezeichnet.

Wie später in Abschn. 2.5 ferner gezeigt wird, liefern der oben berechnete minimale Wert von 2 bit/Bedeutung die von Shannon (1948) eingeführte *informationstheoretische Entropie* bei gleichwahrscheinlichen Bedeutungen und der minimale Wert von $1\frac{3}{4}$ bit pro Bedeutung die *informationstheoretische Entropie* für die Wahrscheinlichkeiten $\frac{1}{2}$, $\frac{1}{4}$, $\frac{1}{8}$, $\frac{1}{8}$. Die informationstheoretische Entropie spielt bei vielen Anwendungen eine zentrale Rolle.

Ein optimaler Codebaum, bei dem jeder Auswahlsschritt zwischen 0 und 1 mit der Wahrscheinlichkeit $\frac{1}{2}$ erfolgt, zeigt in anschaulicher Form auch die *Höhe der Information* an, die dem empfangenden Kommunikationspartner KP_2 mit jeder vom sendenden Kommunika-

tionspartner KP_1 ausgewählten einzelnen Bedeutung geliefert wird. Die Informationshöhe ist gleich der Anzahl der binären Auswahlsschritte im optimalen Codebaum, die benötigt werden, um von der Wurzel (das ist der Punkt W_1) zur jeweiligen Bedeutung zu gelangen. Der Codebaum in Abb. 2.1c ist optimal für $P(S_1) = \frac{1}{2}$, $P(S_2) = \frac{1}{4}$, $P(S_3) = P(S_4) = \frac{1}{8}$. Dort haben also die Bedeutung S_1 die Informationshöhe 1 bit, die Bedeutung S_2 die Informationshöhe 2 bit und die die Bedeutungen S_3 und S_4 je die Informationshöhe 3 bit, was in Abschn. 2.2 noch näher begründet wird. Beim optimalen Codebaum wird der Bedeutungscode KP_1 auch als *Quellencode* bezeichnet. Die Ermittlung des Quellencodes heißt *Quellencodierung*.

Es gibt klassische allgemeine Verfahren für die Konstruktion des optimalen Codebaums, wenn die Anzahl der Bedeutungen und deren Auswahlwahrscheinlichkeiten gegeben sind. Diese Verfahren müssen hier nicht beschrieben werden, man kann sie in vielen Lehrbüchern nachlesen, auch z.B. bei Steinbuch/ Rupprecht (1973).

[2.2] Quellencodierung und Informationshöhe:

Die Quellencodierung besteht in der Ermittlung des optimalen Codebaums auf der Sendeseite (Informationsquelle). Dieser optimale Codebaum hängt nur von den a-posteriori-Wahrscheinlichkeiten ab, mit denen aus Sicht der Empfangsseite einzelne Bedeutungen auf der Sendeseite ausgewählt werden. Der optimale Codebaum liefert mit der Anzahl der Auswahlsschritte (bit) zu einer Bedeutung zugleich auch deren auf der Empfangsseite erhaltene (a-posteriori-) Informationshöhe in bit.

2.1.2 Über weitere Rollen der Codebäume und Bedeutungscode

Was ein Empfänger als Information erfährt, wird auf der Sendeseite erzeugt. Auf der Empfangsseite geschieht der Vorgang des Erfahrens oder Bewusstwerden von Information beim Endnutzer folgendermaßen: Die Decodierung des empfangenen Signals s_i liefert das zugehörige Codewort des Bedeutungscode KP_2 . Dieses Codewort kennzeichnet den Pfad von der Wurzel W_2 zur zugehörigen Bedeutung S_i , es lokalisiert also die zugehörige Bedeutung in der empfangsseitigen Bedeutungsliste. Für die in Abb. 2.1a dargestellte Situation liefert die Decodierung die Beziehungen (2.3).

Wenn der Codebaum in Abb. 2.1a sowohl auf der Sende- als auch auf der Empfangsseite durch den Codebaum in Abb. 2.1c ersetzt wird, dann gelten, wie bereits gesagt, statt (2.1) die Beziehungen (2.5). An die Stellen der Beziehungen (2.2) und (2.3) treten jetzt die folgenden:

$$\left. \begin{array}{l} \text{Sendeseite:} \quad 0 \rightarrow s_1; \quad 10 \rightarrow s_2; \quad 110 \rightarrow s_3; \quad 111 \rightarrow s_4 \\ \text{Empfangsseite:} \quad s_1 \rightarrow 0; \quad s_2 \rightarrow 10; \quad s_3 \rightarrow 110; \quad s_4 \rightarrow 111 \end{array} \right\} \quad (2.7)$$

Wird z.B. das Signal s_2 empfangen, dann liefert die Decodierung das Codewort 10 und dieses wiederum mit (2.5) die Bedeutung S_2 . Diese Bedeutung S_2 wird mit zwei Entscheidungsschritten gefunden. Bei Empfang des Signals s_3 wird die zugehörige Bedeutung S_3 erst nach drei Entscheidungsschritten gefunden. Nach jeder Lokalisierung einer Bedeutung beginnt mit der Decodierung des nächsten Signals die Pfadbildung zur zugehörigen nächsten Bedeutung. Ergänzend hierzu seien noch zwei andersartige Fälle (A) und (B) diskutiert:

A: Ungleiche Codebäume beim Sender und Empfänger

Die bisher betrachtete Verwendung des gleichen Codebaums auf der Sendeseite und der Empfangsseite stellt einen Sonderfall dar, der dadurch gekennzeichnet ist, dass der Bedeutungscode KP_2 die genaue Umkehrung des Bedeutungscode KP_1 darstellt, siehe (2.7). Es ist nämlich keineswegs notwendig, dass auf beiden Seiten der gleiche Codebaum vorhanden ist. In der Natur ist es vielmehr so, dass im Normalfall die Codebäume auf beiden Seiten verschieden sind. Diese Verschiedenheit zeigt sich auch darin, dass auf der Sendeseite und Empfangsseite in der Regel unterschiedlich große Listen an Bedeutungen vorhanden sind.

Bevor auf den Fall unterschiedlich großer Bedeutungslisten eingegangen wird, sei als Beispiel zuerst der einfache Fall betrachtet, dass auf der Sendeseite der in Abb. 2.1b gezeigte Codebaum und auf der Empfangsseite der in Abb. 2.1c gezeigte Codebaum vorhanden ist. In diesem Fall gilt der folgende Satz von Beziehungen:

$$\begin{array}{ll}
 \text{Sendeseite:} & \begin{array}{ll} 00 \rightarrow S_1; & 01 \rightarrow S_2; & 10 \rightarrow S_3; & 11 \rightarrow S_4 & (a) \\ 00 \rightarrow s_1; & 01 \rightarrow s_2; & 10 \rightarrow s_3; & 11 \rightarrow s_4 & (b) \end{array} \\
 \text{Empfangsseite:} & \begin{array}{ll} s_1 \rightarrow 0; & s_2 \rightarrow 10; & s_3 \rightarrow 110; & s_4 \rightarrow 111 & (c) \\ 0 \rightarrow S_1; & 10 \rightarrow S_2; & 110 \rightarrow S_3; & 111 \rightarrow S_4 & (d) \end{array}
 \end{array} \quad \left. \vphantom{\begin{array}{l} (a) \\ (b) \\ (c) \\ (d) \end{array}} \right\} \quad (2.8)$$

Will der sendende Kommunikationspartner KP_1 z.B. die Bedeutung S_3 mitteilen, dann wählt er das Codewort 10 und ordnet diesem das Signal s_3 zu. Dieses Signal s_3 gelangt zum empfangenden Kommunikationspartner KP_2 , der (bei richtiger Interpretation) dieses Signal s_3 als Codewort 110 decodiert und damit über den Codebaum in Abb. 2.1c die auf der Sendeseite gemeinte Bedeutung S_3 richtig findet. Dieses Beispiel macht deutlich, dass es auf die Form der Signale, die der sendende Kommunikationspartner KP_1 zuordnet, gar nicht besonders ankommt. Der empfangende Kommunikationspartner KP_2 muss die Signale nur *richtig interpretieren* können. Die Interpretation eines empfangenen Signals s_i , das sei hier nochmal gesagt, ist ein zweiteiliger Vorgang. Er besteht 1.) in der Decodierung des Signals s_i , die das zugehörige Codewort des Bedeutungscode KP_2 liefert, und 2.) in dem damit ermöglichten Auffinden der zugehörigen Bedeutung.

Eine Besonderheit im Beziehungssatz (2.8) liegt darin, dass die gleichen Bedeutungen S_i durch unterschiedlich viele bit auf der Sendeseite und Empfangsseite gekennzeichnet sind. Auf der Sendeseite kennzeichnet im Fall, dass dort der Codebaum optimal ist, die Anzahl der bit die a-posteriori-Wahrscheinlichkeit, mit der aus Sicht der Empfangsseite die betreffende Bedeutung auf der Sendeseite ausgewählt wird, siehe [2.2]. Kenntnis über die Höhe dieser a-posteriori-Wahrscheinlichkeiten kann der empfangende Kommunikationspartner KP_2 erst im Verlauf von wiederholten und längeren Kommunikationsprozessen gewinnen. Das ist bei der natürlichen Sprache der Fall, von der der Philosoph L. Wittgenstein sagt, dass sie durch den Gebrauch entstehe. Wenn der empfangende Kommunikationspartner KP_2 erstmals mit dem sendenden Kommunikationspartner KP_1 in Kontakt tritt, dann erwartet er von ihm die Auswahl einer bestimmten Bedeutung mit einer a-priori-Wahrscheinlichkeit. Diese a-priori-Wahrscheinlichkeit ist im Allgemeinen von der a-posteriori-Wahrscheinlichkeit verschieden und nur im Sonderfall der Erwartungstreue gleich. Die Höhe der a-priori-Wahrscheinlichkeit hängt mit der Anzahl der bit zusammen, mit der auf der Empfangsseite die einzelnen Bedeutungen S_i gekennzeichnet werden. Das folgt aus einer ähnlichen Überlegung, wie sie am Ende von Abschn. 2.1.1 angestellt wurde.

Einführung in die Theorie der kognitiven Kommunikation
Wie Sprache, Information, Energie, Internet, Gehirn und
Geist zusammenhängen

Rupprecht, W.

2014, XVII, 491 S. 109 Abb., Softcover

ISBN: 978-3-658-05497-7