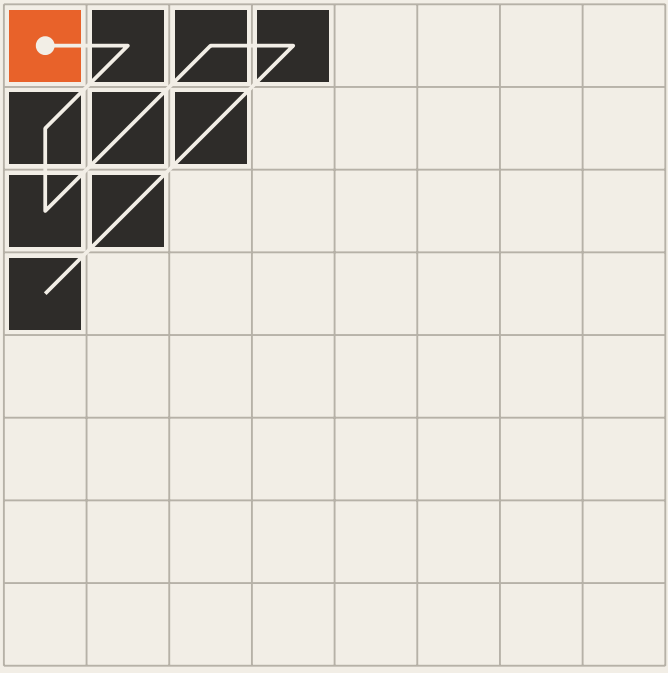


JPEG · ITU-T T.81 · 1992

Die Kompression

Die Kunst des Weglassens

Reibold

Brain-Media.de

Holger Reibold

Die Kompression

Die Kunst des Weglassens

BRAIN-MEDIA.DE

Alle Rechte vorbehalten. Ohne ausdrückliche, schriftliche Genehmigung des Verlags ist es nicht gestattet, das Buch oder Teile daraus in irgendeiner Form durch Fotokopien oder ein anderes Verfahren zu vervielfältigen oder zu verbreiten. Dasselbe gilt auch für das Recht der öffentlichen Wiedergabe. Der Verlag macht darauf aufmerksam, dass die genannten Firmen- und Markennamen sowie Produktbezeichnungen in der Regel marken-, patent- oder warenrechtlichem Schutz unterliegen.

Verlag und Autor übernehmen keine Gewähr für die Funktionsfähigkeit beschriebener Verfahren und Standards.

Das **Umschlagmotiv** zeigt einen Block aus acht mal acht Feldern, wie ihn das JPEG-Verfahren verwendet. Das hervorgehobene Feld links oben enthält die mittlere Helligkeit des Blocks, die Zickzack-Linie die Reihenfolge, in der die Werte gelesen werden.

© 2026 Brain-Media.de

ISBN: 978-3-95444-390-1

Reihe: Morgen war gestern – bahnbrechende Tech-Konzepte, Band 5

Cover: Brain-Media.de

Dr. Holger Reibold – Huber-Müller-Str. 52 – 66113 Saarbrücken

info@brain-media.de – www.brain-media.de

Inhaltsverzeichnis

Vorwort....1
Ein Setzkasten in Morristown....2
Zwei Arten, etwas wegzulassen....3
Worum es in diesem Buch geht....4

1 Weniger Zeichen, gleiche Botschaft....7
1.1 Die Ökonomie der Kürze....7
1.2 Vail, Morse und der Setzkasten....9
1.3 Gerke, Hamburg: der Morsecode wird international....12
1.4 Kurzschrift, Codebücher und andere Vorläufer....13
1.5 1948: Claude Shannon misst die Information....15
1.6 Redundanz: warum Sprache länger ist, als sie sein müsste....18
1.7 Vom Fernschreiber zu UTF-8....21
1.8 Management Summary....23

2 Die Mathematik des Weglassens....25
2.1 Huffman 1951: Hausarbeit statt Prüfung....25
2.2 Der Huffman-Baum: optimal für Zeichen....27
2.3 Haifa 77/78: Lempel und Ziv – das selbstlernende Wörterbuch 29
2.4 Die Familie wächst: LZSS und LZW....32
2.5 Das Bündnis: DEFLATE verbindet LZ77 und Huffman....34
2.6 Ein Blick in eine ZIP-Datei....36
2.7 Alternative Verfahren: Aritmetrik, PPM und Burrows-Wheeler. 37
2.8 Die Grenze: warum sich Zufall nicht komprimieren lässt....39
2.9 Management Summary....42

3 Der Kampf um das Format.....45
3.1 Modems und Faxgeräte: Kompression wird unverzichtbar.....45
3.2 ARC vs. ZIP: Phil Katz verliert Prozess und gewinnt Standard...47
3.3 Stacker, DoubleSpace und der Microsoft-Prozess.....49
3.4 Weihnachten 1994: Unisys verlangt Lizenzgebühren für GIF....51
3.5 PNG: ein freies Format entsteht in zwei Monaten.....52
3.6 gzip, zlib und HTTP: die stille Kompression jeder Webseite.....54
3.7 Die Archivierer: RAR, 7-Zip und LZMA.....55
3.8 Nach DEFLATE: Brotli und Zstandard.....57
3.9 Unsichtbare Kompression in Code und Datenbanken.....58
3.10 Jedes Bit zählt: Kompression im Weltraum.....60
3.11 Management Summary.....62

4 Was das Auge nicht sieht.....65
4.1 Ein Bild ist eine Zahlenwüste.....65
4.2 Helligkeit vor Farbe: wie das Auge Information gewichtet.....66
4.3 1972: Nasir Ahmed und die Kosinustransformation.....68
4.4 1992: JPEG zerlegt jedes Bild in Kacheln.....71
4.5 Lena: das umstrittenste Testbild der Informatik.....74
4.6 JPEG: Wegbereiter für Web, Digicam und Smartphone.....75
4.7 Nachfolger, die keiner nahm.....77
4.8 Oktober 2026: JPEG XL kehrt zurück.....78
4.9 Management Summary.....81

5 Was das Ohr nicht hört.....85
5.1 Die CD und das Problem mit 1,4 Megabit pro Sekunde.....85
5.2 Psychoakustik: Frequenzgruppen und Verdeckung.....87
5.3 Von der Idee zum Verfahren: Erlangen in den 1980er-Jahren...90
5.4 MPEG: drei Schichten und ein Wettbewerb.....92
5.5 Layer II gewinnt den Rundfunk, Layer III das Internet.....94

5.6 Winamp, Napster, iPod: MP3 verändert die Musikindustrie.....95
5.7 Hört man den Unterschied?.....98
5.8 Management Summary.....102

6 Bilder in Bewegung.....105
6.1 Ein Film ist eine Folge fast gleicher Bilder.....105
6.2 Bewegungsschätzung: nur übertragen, was sich verändert....106
6.3 Von H.261 über MPEG-2 zu DVD und Digitalfernsehen.....109
6.4 2003: H.264 und die Rolle des Fraunhofer HHI.....110
6.5 HEVC, VVC und der Rhythmus der Halbierung.....112
6.6 AV1 und der Lizenzkrieg.....114
6.7 Streaming: wenn Video das Internet füllt.....116
6.8 Management Summary.....119

7 Der Preis des Weglassens.....121
7.1 Artefakte: Blöcke, Ringe und verwaschene Kanten.....121
7.2 Generationsverlust: die Kopie der Kopie der Kopie.....123
7.3 Archivieren: was in hundert Jahren noch lesbar sein soll.....125
7.4 Kompression als Sicherheitsrisiko.....126
7.5 Medizin und Forensik: wo Weglassen nicht erlaubt ist.....130
7.6 Weniger Daten, weniger Energie?.....132
7.7 Was Verantwortliche tun können.....134
7.8 Management Summary.....136

8 Verstehen heißt komprimieren.....139
8.1 Vorhersagen und Komprimieren sind dasselbe.....139
8.2 Context Mixing und der Hutter-Preis.....141
8.3 2023: Sprachmodelle schlagen PNG und FLAC.....143
8.4 Ein unscharfes JPEG des Webs?.....145
8.5 Neuronale Codecs: die Rückkehr der Bildkompression.....147

8.6 Wenn der Decoder dazuerfindet: Halluzination statt Artefakt..149
8.7 Drei mögliche Zukünfte.....151
8.8 Management Summary.....154

Zum Schluss.....155

Anhang A: Glossar.....159
Anhang B: Die Geschichte der Kompression in Daten.....167
Anhang C: Die wichtigsten Personen und Organisationen.....173
Anhang D: Die Mathematik des Weglassens.....179
Literatur- und Quellenverzeichnis.....185
Stichwortverzeichnis.....191

Mehr von Brain-Media.de.....195

Vorwort

Nehmen Sie Ihr Smartphone zur Hand und machen Sie ein Foto. Irgendetwas, den Schreibtisch, den Blick aus dem Fenster, eine Tasse Kaffee. Das Bild, das jetzt im Speicher liegt, besteht aus rund zwölf Millionen Bildpunkten. Jeder dieser Punkte hat eine Farbe, und jede Farbe wird durch drei Zahlen beschrieben, eine für Rot, eine für Grün, eine für Blau. Wer das ausrechnet, kommt auf etwa 36 Millionen Zahlen, also rund 36 Megabyte.

Schauen Sie nach, wie groß die Datei tatsächlich ist. Meist sind es zwei bis vier Megabyte. Neun Zehntel der Daten sind verschwunden, und Sie haben nichts davon bemerkt. Das Foto sieht aus wie das, was Sie fotografiert haben.

Mit Musik ist es ähnlich. Ein Lied von vier Minuten Länge belegt auf einer CD gut 40 Megabyte. Als MP3-Datei sind es oft weniger als vier. Bei Filmen werden die Zahlen schwindelerregend. Ein zweistündiger Spielfilm in 4K-Auflösung würde ohne Kompression mehrere Terabyte belegen, also mehrere tausend Gigabyte. Gestreamt wird er mit einem Bruchteil davon, typischerweise mit gut zehn Gigabyte. Der Faktor liegt bei mehreren Hundert.

Ohne Kompression gäbe es kein Streaming, keine Videokonferenzen, keine Fotografie mit dem Telefon, kein Musikhören unterwegs und kein Web, wie wir es kennen. Die Datenleitungen der Welt würden unter der Last zusammenbrechen, die Speicher wären nach wenigen Tagen voll. Kompression ist die unsichtbarste aller Schlüsseltechni-

ken. Sie arbeitet in jedem Gerät, bei jedem Seitenaufruf, bei jedem Anruf, und fast niemand denkt je darüber nach.

Ein Setzkasten in Morristown

Dabei ist Kompression keine Erfindung des Computerzeitalters. Ihre Geschichte beginnt in einer Zeitungsdruckerei in einer Kleinstadt in New Jersey.

Ende der 1830er-Jahre arbeitete dort ein junger Mechaniker namens Alfred Vail mit dem Maler und Erfinder Samuel Morse an einem elektrischen Telegrafen. Das Gerät konnte nur zwei Dinge: Strom an und Strom aus, kurz oder lang. Um damit Buchstaben zu übertragen, brauchte man einen Code, eine Zuordnung von Buchstaben zu Folgen aus kurzen und langen Signalen. Vail fragte sich, welche Buchstaben die kürzesten Folgen bekommen sollten. Die Antwort fand er in den Setzkästen der Druckerei. Die Fächer mit den meisten Lettern, also den Buchstaben, die beim Drucken am häufigsten gebraucht wurden, gehörten dem E und dem T. Das E bekam einen einzigen kurzen Impuls, das T einen einzigen langen. Seltene Buchstaben wie Q oder Z bekamen lange Folgen.

Damit war das Grundprinzip der Kompression gefunden, noch ehe es jemand so nannte: Was häufig vorkommt, wird kurz geschrieben. Hundert Jahre später machte der amerikanische Mathematiker Claude Shannon daraus eine Theorie. Von da an führt eine erstaunlich gerade Linie über die Algorithmen der 1950er- und 1970er-Jahre bis zu den Dateiformaten, die heute jeden Bildschirm füllen.^

Zwei Arten, etwas wegzulassen

Wer etwas komprimiert, lässt etwas weg. Die ganze Geschichte dieses Buches ist die Geschichte einer einzigen Frage: Was darf man weglassen?

Auf diese Frage gibt es zwei grundverschiedene Antworten. Die erste lautet: das, was sich wiederholt. Ein Text enthält viele gleiche Wörter, eine Tabelle viele gleiche Werte, ein Programmcode viele gleiche Befehle. Wer jede Wiederholung durch einen kurzen Verweis auf die erste Stelle ersetzt, spart Platz, und er kann das Original später Bit für Bit wiederherstellen. Diese verlustfreie Kompression steckt in jeder ZIP-Datei, in jedem PNG-Bild und in fast jeder Webseite, die Sie aufrufen. Ein Vertrag, eine Datenbank oder ein Programm dürfen auch nicht um ein einziges Zeichen verändert werden. Für sie ist sie die einzige Möglichkeit.

Die zweite Antwort ist kühner: das, was niemand bemerkt. Unsere Augen und Ohren sind keine Messgeräte. Sie nehmen manche Unterschiede sehr genau wahr, andere gar nicht. Ein leiser Ton neben einem lauten verschwindet für das Ohr. Feine Farbabstufungen in einer unruhigen Fläche sieht das Auge nicht. Wer genau weiß, was Menschen wahrnehmen, kann alles andere fortwerfen. Diese verlustbehaftete Kompression stellt das Original nie wieder her. Sie liefert etwas, das sich gleich anfühlt. Auf ihr beruhen JPEG, MP3 und alle Verfahren, mit denen Filme übertragen werden.

Beide Antworten haben ihre Helden, ihre Irrtümer und ihre Kämpfe. Es gibt Patentstreitigkeiten, die ganze Formate zu Fall brachten, und Prozesse, die Verlierer zu Gewinnern machten. Es gibt einen Studenten, der eine Hausarbeit schrieb, statt eine Prüfung abzulegen, und

damit seinen Professor übertraf. Und es gibt eine Gruppe von Ingenieuren in Franken, die etwas für unmöglich Gehaltenes möglich machten und dann zusehen mussten, wie ihre Erfindung ohne sie die Welt eroberte.

Worum es in diesem Buch geht

Wie die anderen Bände dieser Reihe verfolgt auch dieser eine einzige Idee. Im Band über den Memristor war es die leere vierte Ecke in einem Diagramm, im Band über den Transformer das Nadelöhr, durch das alle Information hindurchmusste. Hier ist es das Weglassen.

Jedes Kapitel endet mit einem Satz darüber, was in der jeweiligen Epoche weggelassen wurde. Am Anfang ist es nur das Überflüssige. Am Ende ist es eine Frage, die weit über die Technik hinausreicht. Denn es zeigt sich, dass gute Kompression und Intelligenz enger verwandt sind, als man lange dachte. Wer einen Text gut zusammenpressen will, muss vorhersagen können, was als Nächstes kommt. Genau das tun die großen Sprachmodelle, von denen der Band über den Transformer erzählt.

Das erste Kapitel führt von der Telegrafie zu Claude Shannon und seiner Theorie der Information. Das zweite erzählt, wie Mathematiker das Weglassen berechenbar machten, von David Huffman über Abraham Lempel und Jacob Ziv bis zu DEFLATE, dem Verfahren, das bis heute in ZIP-Dateien und PNG-Bildern arbeitet. Das dritte Kapitel handelt von den Formatkriegen der 1980er- und 1990er-Jahre und von der Kompression, die heute unbemerkt in Browsern, Datenbanken und Softwarepaketen läuft. Im vierten Kapitel geht es um Bilder und um JPEG, im fünften um Musik und MP3, im sechsten um Film und Video.

Das siebte Kapitel fragt, was das Weglassen kostet: an Qualität, an Sicherheit und an Gewissheit. Das achte schließlich zeigt, warum die Kompression heute in der künstlichen Intelligenz eine überraschende Rückkehr erlebt.

Zwischen den Kapiteln stehen kurze Zwischenstücke, die Erlanger Intermezzi. Sie erzählen die Geschichte der MP3 von einer Idee aus den 1970er-Jahren bis zum Ende des Lizenzprogramms im Jahr 2017. Diese Geschichte läuft unter dem ganzen Buch mit, auch dort, wo die Kapitel von Bildern, Texten oder Videos handeln, und sie kommt in Kapitel 5 an die Oberfläche.

Jedes Kapitel endet zudem mit einem Management Summary, einer kurzen Zusammenfassung für eilige Leser. Wer die Mathematik hinter den Verfahren kennenlernen möchte, findet sie in Anhang D. Für das Verständnis des Haupttextes ist sie nicht nötig. Und wer selbst ausprobieren will, wovon die Rede ist, findet in den Kapiteln immer wieder kleine Versuche, die sich mit jedem Computer durchführen lassen.

Herzlichst

Holger Reibold

Erlanger Intermezzo I

Ein Professor und das Telefonnetz

Erlangen, Ende der 1970er-Jahre. Dieter Seitzer ist Professor für Technische Elektronik an der Universität. Er hat eine Idee, die seine Kollegen für versponnen halten.

Die Deutsche Bundespost plant ein neues, digitales Telefonnetz. Es soll später ISDN heißen. Seitzer fragt sich, ob man über diese Leitungen nicht nur Gespräche, sondern auch Musik schicken könnte, in einer Qualität, die der einer Schallplatte nahekommt.

Die Rechnung spricht dagegen. Musik in CD-Qualität braucht mehr als zehnmal so viel Bandbreite, wie eine ISDN-Leitung bietet. Man müsste also mehr als neun Zehntel der Daten weglassen, ohne dass es jemand hört.

Seitzer hält das für machbar. Er kennt die Forschung zur Psychoakustik, zur Frage, was das menschliche Ohr überhaupt wahrnimmt. Nach seiner eigenen Darstellung reicht er sogar einen Patentantrag für eine Art digitale Musikbox ein. Das Patentamt hält die Idee für nicht ausführbar.

Er ist nicht der Einzige, der so denkt. An der Technischen Universität München hat Eberhard Zwicker in jahrzehntelanger Arbeit vermessen, wie das Ohr Töne verarbeitet und was es überhört. Seitzer kennt diese Arbeiten. Er ahnt, dass in ihnen der Schlüssel liegt.

Seitzer braucht jemanden, der es trotzdem versucht.

1 Weniger Zeichen, gleiche Botschaft

Versuchen Sie, diesen Satz zu lesen: Ds lsst sch ch hn Vkl lsn.

Sie haben vermutlich einen Moment gebraucht, aber Sie haben es geschafft. Der Satz lautet: Das lässt sich auch ohne Vokale lesen. Ein Drittel der Buchstaben fehlt, und die Botschaft ist trotzdem angekommen. Ihr Gehirn hat ergänzt, was fehlte, weil die Sprache mehr Zeichen enthält, als sie eigentlich braucht.

Diese Eigenschaft der Sprache hat einen Namen. Man nennt sie Redundanz, Überfluss. Sie ist der Rohstoff jeder Kompression. Dieses Kapitel erzählt, wie Menschen lange vor dem Computer gelernt haben, diesen Überfluss abzuschöpfen, und wie ein Mathematiker in den 1940er-Jahren herausfand, wie groß er eigentlich ist.

1.1 Die Ökonomie der Kürze

Im 19. Jahrhundert war Kürze bares Geld. Wer ein Telegramm aufgab, bezahlte nach der Zahl der Wörter. Eine Nachricht über den Atlantik kostete nach der Eröffnung der ersten dauerhaft funktionierenden Kabelverbindung im Jahr 1866 ein kleines Vermögen. Kaufleute, Reedereien und Banken hatten deshalb ein starkes Interesse daran, möglichst viel Inhalt in möglichst wenige Wörter zu packen.

Auch wer keine Codebücher besaß, lernte zu sparen. Der Telegrammstil, der bis heute als Inbegriff knapper Sprache gilt, ist ein Kind der Wortgebühr. Artikel, Hilfsverben und Höflichkeiten fielen weg, übrig blieb das Gerüst der Botschaft: Ankomme Dienstag Zug zwölf Uhr bitte

abholen. Weil Satzzeichen ebenfalls kosteten oder missverstanden werden konnten, schrieb man statt eines Punktes das Wort „stop“. Jeder Empfänger verstand, was gemeint war, weil er die fehlenden Wörter aus dem Zusammenhang ergänzte. Der Telegrammstil ist damit eine verlustbehaftete Kompression im Kopf des Absenders, die sich auf das Weltwissen des Empfängers verlässt.

Daraus entstand eine eigene Literaturgattung, die Telegrafencodebücher. Sie enthielten Tausende vorformulierter Sätze, von denen jeder ein einzelnes Codewort zugeordnet bekam. Ein Kaufmann, der seinem Geschäftspartner mitteilen wollte, dass die Ware eingetroffen sei, aber in beschädigtem Zustand, und dass er eine Preisminderung erwarte, schrieb dafür ein einziges künstliches Wort. Der Empfänger schlug es im selben Codebuch nach und las den ganzen Satz. Bekannte Werke wie der ABC-Code aus den 1870er-Jahren oder der Bentley-Code von 1906 wurden in hohen Auflagen gedruckt und in Kontoren in aller Welt benutzt.

Man kann diese Codebücher als frühe Kompressionsverfahren betrachten. Sie ersetzten häufige, lange Botschaften durch kurze Zeichen. Sie hatten allerdings einen Haken, der sich durch die ganze Geschichte der Kompression ziehen wird: Sender und Empfänger brauchten dasselbe Wörterbuch. Ohne das Buch war die Nachricht wertlos. Und das Buch funktionierte nur für die Sätze, die jemand vorher hineingeschrieben hatte. Für alles Unvorhergesehene musste man wieder im Klartext telegrafieren und Wort für Wort bezahlen.

Eine verwandte Form der Verdichtung hat bis heute überlebt, die Q-Gruppen des Funkverkehrs. Seit den internationalen Funkkonferenzen zu Beginn des 20. Jahrhunderts stehen Kürzel aus drei Buchstaben, die alle mit Q beginnen, für ganze Fragen und Antworten. QTH bedeutet:

Mein Standort ist. QRZ fragt: Wer ruft mich? Ein Schiffsfunker in der Nordsee und einer im Mittelmeer konnten sich so verständigen, ohne eine gemeinsame Sprache zu sprechen. Funkamateure benutzen diese Kürzel bis heute. Auch hier steckt das Prinzip im Hintergrund: ein vereinbartes Wörterbuch, kurze Zeichen für häufige Botschaften, und alles, was nicht im Wörterbuch steht, muss ausgeschrieben werden.

Die Telegrafenverwaltungen sahen das Treiben mit gemischten Gefühlen. Sie verdienten am Wort, und die Codebücher schmälerten ihre Einnahmen. Auf internationalen Konferenzen wurde deshalb immer wieder darüber gestritten, wie lang ein Codewort sein durfte und ob künstliche Wörter überhaupt als Wörter zählten. Kompression war von Anfang an nicht nur eine technische, sondern auch eine wirtschaftliche Frage: Wer spart, wenn weniger übertragen wird, und wer verliert?

1.2 Vail, Morse und der Setzkasten

Noch grundlegender als die Codebücher war der Code, auf dem die Telegrafie selbst beruhte. Seine Geschichte beginnt nicht mit dem Mann, nach dem er benannt ist.

Samuel Morse war in den 1830er-Jahren ein angesehener Porträtmaler, der sich für Elektrizität zu begeistern begann. Er entwarf einen Telegrafen, der Signale über einen Draht schickte und am anderen Ende auf einem Papierstreifen aufzeichnete. Für die Übertragung von Nachrichten dachte er zunächst an ein umständliches Verfahren: Jedes Wort sollte eine Nummer bekommen, die Nummern sollten übertragen und am Ziel in einem dicken Wörterbuch nachgeschlagen werden. Das war im Grunde ein Codebuch, mit allen Nachteilen, die solche Bücher hatten.

Der Anstoß dazu soll ein persönlicher Verlust gewesen sein. Nach einer verbreiteten Darstellung erhielt Morse 1825 während eines Porträtauftrags in Washington die Nachricht, dass seine Frau schwer erkrankt sei. Als er nach Hause kam, war sie bereits gestorben und beerdigt. Die Briefpost war zu langsam gewesen. Ob dieses Erlebnis tatsächlich den Ausschlag gab, lässt sich nicht sicher sagen. Fest steht, dass Morse in den folgenden Jahren davon besessen war, Nachrichten schneller als ein Pferd zu befördern.

Alfred Vail, der Sohn eines Eisenwerkbesitzers aus Morristown in New Jersey, sah 1837 eine Vorführung von Morses Apparat und wurde sein Partner. Er brachte Geld, handwerkliches Geschick und eine bessere Idee mit. Statt ganzer Wörter sollten einzelne Buchstaben übertragen werden, jeder durch eine kurze Folge aus kurzen und langen Signalen, Punkten und Strichen. Damit ließ sich jedes beliebige Wort senden, ohne Wörterbuch.

Blieb die Frage, welcher Buchstabe welche Folge bekommen sollte. Vail erkannte, dass es nicht gleichgültig war. Jeder Punkt und jeder Strich kostete Zeit, und Zeit auf der Leitung war Geld. Wenn häufige Buchstaben kurze Folgen bekamen, wurden die Nachrichten insgesamt kürzer. Nur: Wie häufig war welcher Buchstabe?

Statistiken über Buchstabenhäufigkeiten gab es damals nicht. Vail fand eine Abkürzung. Er ging in die Druckerei der örtlichen Zeitung und zählte die Lettern in den Setzkästen. Ein Drucker braucht von jedem Buchstaben so viele Exemplare, wie er auf einer Seite benötigt. Die Setzkästen waren also eine Häufigkeitstabelle der englischen Sprache aus Blei. Das E lag dort in größter Menge, gefolgt vom T, vom A und vom O. Das Q, das X und das Z gab es nur in wenigen Exemplaren.

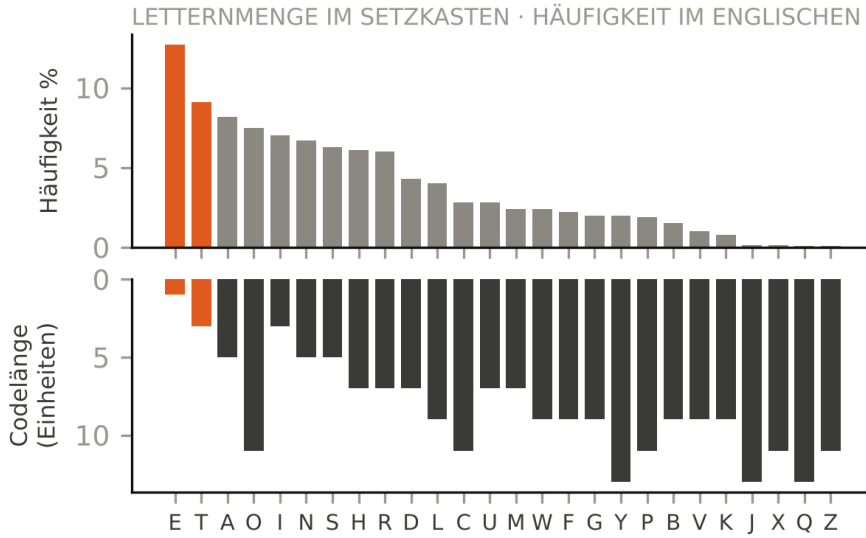


Abbildung 1: Der Setzkasten als Häufigkeitstabelle. Je mehr Lettern ein Buchstabe im Kasten hatte, desto kürzer wurde sein Code. Dargestellt sind die Häufigkeit im Englischen und die Länge des Morsecodes in Signaleinheiten.

So bekam das E einen einzigen Punkt und das T einen einzigen Strich. Häufige Buchstaben wie A, I, N und M erhielten Folgen aus zwei Zeichen, seltene wie Q, Y oder J lange Folgen aus vier Zeichen. Am 24. Mai 1844 ging über die erste Versuchsleitung zwischen Washington und Baltimore die berühmte Nachricht „What hath God wrought“. Der Code, mit dem sie übertragen wurde, trug Morses Namen. Vail, der ihn ausgearbeitet hatte, erhielt wenig Anerkennung und verließ die Telegrafie einige Jahre später enttäuscht.

Vails Idee war schlicht, und sie ist bis heute der Kern vieler Kompressionsverfahren. Wenn die Zeichen einer Nachricht nicht gleich häufig sind, lohnt es sich, ihnen unterschiedlich lange Codes zu geben. Die kurzen Codes für die häufigen Zeichen sparen mehr ein, als die langen Codes für die seltenen kosten. Ein Jahrhundert später sollte ein Stu-

dent am Massachusetts Institute of Technology zeigen, wie man diese Zuordnung nicht nach Gefühl, sondern beweisbar optimal vornimmt.

1.3 Gerke, Hamburg: der Morsecode wird international

Der Code, den Funker und Pfadfinder heute als Morsealphabet lernen, ist nicht mehr der von Vail. Er stammt zu einem guten Teil aus Hamburg.

Friedrich Clemens Gerke, 1801 geboren, war Schriftsteller, Journalist und Telegrafenpionier. Als 1848 die Elektro-Magnetische Telegraphen-Compagnie zwischen Hamburg und Cuxhaven eine der ersten elektrischen Telegrafenlinien Europas in Betrieb nahm, wurde er ihr Inspektor. Der regelmäßige Betrieb begann am 15. Juli 1848. Die Linie diente vor allem der Schifffahrt. Sie meldete einlaufende Schiffe von der Elbmündung an die Kaufleute in der Stadt.

Gerke übernahm den amerikanischen Code, stieß aber schnell auf dessen Schwächen. Vails Code kannte nicht nur Punkte und Striche, sondern mehrere verschieden lange Striche und innerhalb mancher Buchstaben zusätzliche Pausen. Das C etwa bestand aus zwei Punkten, einer kurzen Pause und einem weiteren Punkt. Auf langen, störanfälligen Leitungen und bei müden Telegrafisten führte das zu Verwechslungen. Eine etwas zu lange Pause verwandelte einen Buchstaben in zwei andere.

Gerke überarbeitete den Code gründlich und veränderte dabei fast die Hälfte der Zeichen. In seiner Fassung gab es nur noch zwei Elemente, den kurzen Punkt und den Strich, der dreimal so lang war. Die Pausen

innerhalb eines Buchstabens waren immer gleich lang. Der Deutsch-Österreichische Telegraphenverein übernahm Gerkes Code 1851. Im Jahr 1865 machte ihn die neu gegründete Internationale Telegraphenunion in leicht veränderter Form zum internationalen Standard. Der amerikanische Code hielt sich noch lange auf den Leitungen der amerikanischen Eisenbahnen, verschwand dann aber.

Gerkes Reform ist für die Geschichte der Kompression lehrreich. Sein Code war im Durchschnitt etwas länger als der amerikanische. Er war weniger stark komprimiert. Dafür war er robuster und leichter zu lernen. Gerke hatte erkannt, was Ingenieure bis heute immer wieder neu entdecken: Die kürzeste Darstellung ist nicht immer die beste. Wer Redundanz entfernt, macht eine Nachricht anfälliger für Störungen. Ein kleiner Fehler in einem stark komprimierten Signal kann großen Schaden anrichten. Davon wird in diesem Buch noch mehrfach die Rede sein.

1.4 Kurzschrift, Codebücher und andere Vorläufer

Die Telegrafie war nicht der einzige Bereich, in dem Menschen lernten, mit weniger Zeichen auszukommen. Schon in der Antike gab es Kurzschriften. Marcus Tullius Tiro, der Sekretär Ciceros, entwickelte ein System, mit dem sich Reden im Senat mitschreiben ließen. Im 19. Jahrhundert verbreitete sich im deutschen Sprachraum die Kurzschrift von Franz Xaver Gabelsberger, aus der später die Deutsche Einheitskurzschrift hervorging. Sie war in Parlamenten, Gerichten und Büros bis weit ins 20. Jahrhundert unentbehrlich.

Mehr von Brain-Media.de

Morgen war gestern. Bahnbrechende Tech-Konzepte

Die Reihe erzählt die Geschichte der Ideen, auf denen die digitale Welt beruht. Jeder Band verfolgt ein einziges Konzept von seinen Anfängen bis in die Gegenwart und ist für sich lesbar.

Band 1: Der Memristor. Das fehlende Bauteil. Wie Computer lernen, sich zu erinnern. Wie ein 1971 vorhergesagtes Bauelement mit Gedächtnis die Architektur des Computers verändern und den Energiehunger der künstlichen Intelligenz zähmen könnte.

Band 2: Der Transformer.

Band 3: Rechenzentren im All.

Band 4: Das Frequenzspringen.

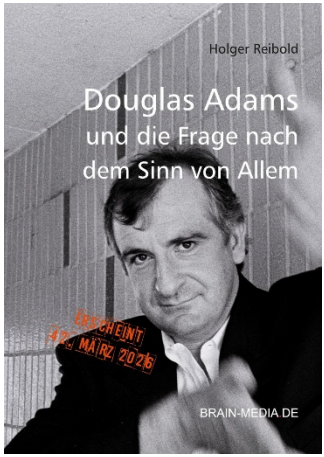
Band 5: Die Kompression.

Brain-Media Compliance Complete

Der Jahreszugang zur Fachbibliothek von Brain-Media.de mit Büchern und Arbeitshilfen zu NIS-2, DORA, Cyber Resilience Act, EU AI Act und ISO 27001. Für alle, die künstliche Intelligenz nicht nur verstehen, sondern auch rechtssicher einsetzen wollen.

Details finden Sie unter folgender URL:

www.brain-media.de/compliance-complete.html



42 – Douglas Adams und die Frage nach dem Sinn von Allem

Am 11. Mai 2026 ist Douglas Adams 25 Jahre tot. Der Kultautor hat der Welt wunderbar, skurrile Werke geschenkt. Jetzt ist es an der Zeit, den Autor kennenzulernen.

Umfang: 140 Seiten

Preis: 14,99 EUR

Erschienen: 42. März 2026



Das Douglas-Adamds-Bundle:

das **Buch** Zweiundvierzig**,**

das **Handtuch** zum Towelday und

das **Hörbuch** für unterwegs.

Mit dem Towelday, erinnern sich Fans an Douglas Adams und huldigen dem Kultautor.

100 % intergalaktisch geprüfte Baumwolle, nachhaltig Produktion.

Exklusiv bei Brain-Media.de: Bundle zum Preis von 42 EUR.