98. Information optimal verpackt

ISBN 0-19-851187-6
9 780198 511878

Die Notwendigkeit, Informationen an Leute weiterzugeben, die sich für das gleiche Gebiet interessieren, taucht in den verschiedensten Lebensbereichen auf. Wie verständigen sich Jazzmusiker über die Harmonien, über die improvisiert werden soll? Wie tanzt man den Grundschritt beim Tango? Welche Nummer gehört zu dem Artikel, den man gerade an der Kasse bezahlen möchte? In der Mathematik hat sich aus dieser Fragestellung ein eigenes Teilgebiet entwickelt, die Codierungstheorie. Genauer geht es dabei darum, Information „optimal“ zu verpacken, wobei die Antwort auf das, was „optimal“ bedeuten soll, je nach Situation unterschiedlich ausfallen kann.

Wie kann man zum Beispiel eine Nachricht so versenden, dass sie vom Empfänger auch dann gelesen werden kann, wenn es bei der Übertragung zu einigen Übermittlungsfehlern kommt? Eine sehr naive Lösung des Problems könnte darin bestehen, die gleiche Nachricht mehrfach – etwa fünf Mal – zu versenden. Es ist dann extrem unwahrscheinlich, dass alle fünf Sendungen an der gleichen Stelle unleserlich oder verfälscht sind, und deswegen kann der Empfänger sicher das, was gemeint ist, herauslesen.

Der Nachteil dieses Verfahrens ist allerdings, dass es extrem unökonomisch ist. Man kann es viel eleganter machen. Dazu wird – wie immer beim Arbeiten mit Computern – jedes der zu übertragenden Zeichen in eine Folge aus Nullen und Einsen übersetzt. Wir nehmen einmal an, dass zu jedem Zeichen eine derartige Folge der Länge 10 gehört. Als einfachen Test, ob sie richtig übermittelt wurde, können wir hinten eine Null oder Eins anhängen, je nachdem, ob die Anzahl der Einsen in der Zehnerfolge gerade war oder nicht. Der Empfänger weiß dann, dass etwas nicht stimmen kann, wenn die „Prüfziffer“ nicht zur Nachricht passt. Mit nur 10 Prozent Mehraufwand kann damit die Richtigkeit der Übertragung kontrolliert werden. Im Fall einer Fehlermeldung ist allerdings noch nicht klar, wo genau der Fehler liegt und – wenn es einen gibt – wie er zu korrigieren ist. All das lässt sich aber mit verfeinerten Verfahren auch noch erreichen.

Die hohe Kunst der Codierungstheorie besteht darin, Nachrichten auch über sehr stark fehlerbehaftete Übertragungskanäle zuverlässig zu schicken. Die Ergebnisse können sich im wahrsten Sinn des Wortes sehen lassen: Ein typisches Beispiel für erfolgreiche Codierungstheorie sind die Bilder, die von weit entfernten Weltraumstationen – etwa vom Mars – gesendet werden. In viel bescheidenerem Rahmen sind derartige Verfahren aber auch in Ihrem CD-Player verwirklicht. Nur aufgrund der eingebauten Codierungstheorie können Sie Ihre Lieblings-CD auch dann störungsfrei abspielen, wenn sie versehentlich einen dicken Kratzer abbekommen hat.

Fehlerkorrigierende Codes

Durch ein Kontrollzeichen kann nur festgestellt werden, dass irgendwo etwas nicht stimmt. Wenn man etwa bei der vorstehend beschriebenen Methode die Nachricht 01100001011 erhält, muss bei der Übertragung etwas schief gegangen sein. Die Anzahl der Einsen ist nämlich ungerade, in allen 11-er Paketen sollte diese Zahl aber gerade sein.

Solche einfachen Tests reichen manchmal aus, z.B. an der Supermarktkasse. Da merkt das Lesegerät, dass es beim Ablesen des Strichcodes eine Panne gegeben hat, und die Kassiererin muss die Ware noch einmal durchziehen.

Manchmal möchte man aber genauer wissen, welches Bit falsch übertragen wurde. Dann kann man es nämlich korrigieren und die ursprüngliche Nachricht rekonstruieren.

Der erste leicht anwendbare derartige Code wurde von *R. W. Hamming* im Jahr 1948 beschrieben, die Idee der Selbstkorrektur war wirklich revolutionär.

Um Hammings Idee zu beschreiben, betrachten wir eine aus vier Zeichen bestehende 0-1-Folge. Wir wollen sie allgemein $a_1a_2a_3a_4$ nennen; geht es etwa um 0110, so wäre $a_1 = 0, a_2 = 1, a_3 = 1, a_4 = 0$. Es sollen nun *drei Kontrollbits angehängt werden*, sie sollen a_4, a_5, a_6 heißen. Die Vorschrift ist die folgende:

- Ist die Anzahl der Einsen in a_1, a_2, a_4 ungerade, so soll $a_5 = 1$ sein; andernfalls wird $a_5 = 0$ gesetzt.
- Ist die Anzahl der Einsen in a_1, a_3, a_4 ungerade, so soll $a_6 = 1$ sein; andernfalls wird $a_6 = 0$ gesetzt.
- Ist die Anzahl der Einsen in a_2, a_3, a_4 ungerade, so soll $a_7 = 1$ sein; andernfalls wird $a_7 = 0$ gesetzt.

Dann wird die aus 7 Zeichen bestehende Folge verschickt: $a_1a_2a_3a_4a_5a_6a_7$. In unserem Beispiel – zu verschicken ist 0110 – wird die Folge vorher zu 0110110 erweitert. Die letzte Null etwa kam deswegen zustande, weil sich unter den Zahlen a_2, a_3, a_4 (also unter den Zahlen $1, 1, 0$) eine gerade Anzahl von Einsen befindet.

Und wozu? Mal angenommen, bei der Übertragung gibt es einen Fehler, eines der Bits wird falsch empfangen. Dann überlegt man sich, dass in den Folgen $a_1a_2a_4a_5$, $a_1a_3a_4a_6$ und $a_2a_3a_4a_7$ jeweils eine gerade Anzahl von Einsen sein müsste, wenn alles fehlerfrei angekommen wäre. Falls nun a_1 falsch empfangen wurde, so hätten $a_1a_2a_4a_5$ und $a_1a_3a_4a_6$ eine ungerade Anzahl von Einsen, und bei $a_2a_3a_4a_7$ wäre alles in Ordnung. Das Muster wäre also ungerade, ungerade, gerade. Wie sieht es bei der fehlerhaften Übertragung anderer Bits aus?

- a_2 fehlerhaft: ungerade, gerade, ungerade.
- a_3 fehlerhaft: gerade, ungerade, ungerade.
- a_4 fehlerhaft: ungerade, ungerade, ungerade.

- a_5 fehlerhaft: ungerade, gerade, gerade.
- a_6 fehlerhaft: gerade, ungerade,gerade.
- a_7 fehlerhaft: gerade, gerade, ungerade.

Man kann also aus der Analyse, wo in $a_1a_2a_4a_5$, $a_1a_3a_4a_6$ und $a_2a_3a_4a_7$ die falschen Einsen-Anzahlen zu finden sind, eindeutig auf das fehlerhafte Bit schließen, es korrigieren und so die Originalfolge in jedem Fall erhalten.

> Ein Beispiel: Mal angenommen, die Folge 0110110 wäre unterwegs zu 1110110 verfälscht worden. Wir betrachten die Einsen-Anzahlen in $a_1a_2a_4a_5$, $a_1a_3a_4a_6$ und $a_2a_3a_4a_7$, also in 1101, 1101 und 1100. Die Einsen-Anzahlen sind ungerade, ungerade bzw. gerade. Der Fehler muss also an der ersten Stelle passiert sein.

Dadurch kann sogar erkannt werden, ob eines der eigentlich uninteressanten Bits a_5, a_6, a_7 falsch ankam. Das ist dann nur ein Indiz dafür, dass der Übertragungskanal nicht ganz zuverlässig ist.

Durch diesen Hamming-Code kann nicht festgestellt werden, ob mehr als ein Fehler vorliegt. Durch ausgefeiltere Codierung können aber auch zwei, drei oder noch mehr Irrtümer in beliebig langen 0-1-Folgen repariert werden. Das ist ganz fundamental für das Funktionieren einer CD, denn ein hundertprozentig makelloses Exemplar wäre schlicht unbezahlbar.

99. Vier Farben reichen immer

Nimmt man ein Blatt Papier und skizziert darauf eine Landkarte, so kann man, um es etwas netter aussehen zu lassen, doch versuchen, die einzelnen Länder zu färben. Dabei wird man sicher aus Gründen der Übersichtlichkeit Länder verschieden färben wollen, die eine gemeinsame Grenze haben.

Wie viele Farben braucht man dafür? Sicher hat man keine Probleme, wenn man so viele Farben hat, wie Länder auf der Karte sind. Es reichen aber offensichtlich weniger, da ja nicht alle Länder mit allen anderen eine gemeinsame Grenze haben können. Überraschenderweise wird man bald feststellen, dass es mit vier Farben immer klappt, wenn man sich nur geschickt genug anstellt.

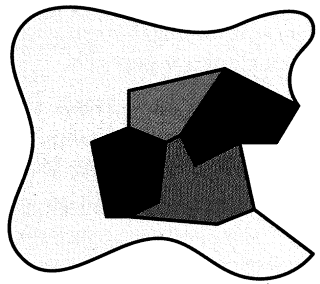

Abbildung 87: Hier braucht man vier Farben (drei reichen nicht)

Das wurde schon im 19. Jahrhundert bemerkt. Mathematiker sind aber mit einem experimentellen Befund nie zufrieden. Es wurde nach einem Beweis gesucht, der garantiert, dass es immer geht: Auch mit noch so vielen und noch so kompliziert ineinander verschachtelten Ländern.

Es wurden große Anstrengungen unternommen, aber das Problem erwies sich als wirklich harte Nuss. Die mathematische Welt musste noch bis in die siebziger Jahre des 20. Jahrhunderts warten, bis das Problem entschieden war: Vier Farben reichen wirklich immer.

Allerdings gibt es mit dem Beweis ein kleines Problem, und deswegen ist das Ganze immer noch in der Diskussion. Er wird zwar nicht angezweifelt, aber wichtige Teile der Argumentation sind an einen Computer delegiert worden. Kein Mensch hätte das bewältigen können, die zugehörigen Rechnungen sind viel zu komplex. Das ist für Mathematiker eine neue und unbefriedigende Situation, an die sie sich (noch) nicht gewöhnen wollen. Auch wenn mehrere Dutzend Computer die Richtigkeit bestätigen, das eigene Aha-Erlebnis hat einen viel höheren Stellenwert.

Für diejenigen, die eher an praktischen Nutzanwendungen interessiert sind, ist das Problem sicher nicht so interessant. Schließlich hält jedes Zeichenprogramm fast beliebig viele Farben bereit. Es gehört eher in die Abteilung „Es gibt unendlich