

Zur Mathematik des Strichcodes – – und einige weitere Bemerkungen

Prof. em. Urs Stammbach
Emeritenstamm, 26. Oktober 2015



Wir alle stammen aus der guten, alten Zeit, aus der Zeit als es noch keine Strichcodes gab. Da war alles ganz anders. Im Dorfladen an der Ecke, in dem wir als Kinder Einkäufe erledigen mussten, stand eine grosse Registrierkasse auf dem Ladentisch. Da tippte die Verkäuferin die Preise ein und nannte am Ende den Totalbetrag.



Heute läuft das alles anders ab. Der Laden heisst heute denn auch neudeutsch Shopping Center. Da findet sich *keine* Registrierkasse der alten Art mehr, und die Person an der Kasse tippt *nicht* die einzelnen Beträge ein. Sie zieht statt dessen einfach die Strichcode-Etiketten der einzelnen Artikel über einen Sensor; der Strichcode wird automatisch gelesen und damit das Produkt identifiziert. Der Computer, der mit dem Lesegerät verbunden ist, addiert die entsprechenden Preise und so nebenbei wird dabei gleich auch die Lagerhaltung organisiert.



Die Strichcode-Etiketten sind – wie Sie wissen – auf jedem Produkt des Detailhandels zu finden. Natürlich haben Sie diese Etiketten schon genauer angesehen und versucht, hinter die Geheimnisse dieser schmalen und breiten Striche zu kommen. Wahrscheinlich hatten Sie dabei keinen durchschlagenden Erfolg, – es sei denn Sie haben sich einigermaßen ernsthaft damit beschäftigt und haben sich im Internet kundig gemacht.

Eigentlich ist es ja störend, dass man über ein Ding, das im täglichen Leben so häufig vorkommt, nicht besser Bescheid weiss! Das sollte man ändern! Aus diesem Grund - sozusagen zur Weiterbildung - habe ich diesen Vortrag vorgeschlagen!

Übrigens: das Lesegerät das an der Ladenkasse verwendet wird, muss ziemlich leistungsfähig sein. Es kann die Etikette weit schneller lesen, als wir das mit unseren Augen schaffen würden, auch wenn wir grosse, deutliche Ziffern vor uns hätten. Und ferner macht dieses Lesegerät offenbar keine Fehler! Oder haben Sie schon einmal erlebt, dass auf Ihrem Kassenzettel ein falsches Produkt aufgeführt worden ist? Eine Flasche Champagner etwa statt einer Flasche Bier? Kaum! Schnell und traumhaft sicher sind diese Geräte! Den Menschen weit überlegen!

Noch eine weitere Bemerkung: Wir sind heute natürlich von anderen Anwendungen mit Lesegeräten verwöhnt, wir benützen Scanner, die gewöhnlich Ziffern erkennen und lesen können, wir kennen Lesegeräte, die Gedrucktes, ja sogar Handschriftliches lesen können, usw. Aber man soll im Zusammenhang mit dem Strichcode bedenken, dass er bereits in den 70er Jahren eingeführt worden ist. Die Technik hat natürlich in der Zwischenzeit grosse Fortschritte gemacht. Aber der Strichcode ist immer noch da und erfüllt ganz offensichtlich die ihm zugeordnete Rolle nach über vierzig Jahren immer noch zur vollen Zufriedenheit der Anwender.

Zur Geschichte

Der Strichcode, von dem wir sprechen, wird auch EAN-Code genannt; EAN ist eine Abkürzung für **E**uropean **A**rticle **N**umbering **S**ystem. (Neuerdings auch **G**lobal **T**rade **I**tem **N**umber, GTIN.) Das System wurde 1977 auf der Grundlage eines kurz vorher in Amerika verwendeten Bezeichnungssystems in Europa eingeführt. Dabei hat man das amerikanische System erweitert, und zwar so, dass die beiden Systeme kompatibel sind.¹

Betrachten wir diesen Strichcode genauer. Man vermutet sofort, dass das Strichmuster in irgend einer Weise die 13 Ziffern darstellt, die unterhalb der Striche zu finden sind.² Das ist richtig, aber die Einzelheiten sind nicht offensichtlich.

Es sind zwei Grundideen, die beim Strichcode zusammenkommen; die eine Idee besteht darin, die Ziffern durch Striche darzustellen; die andere Idee ist etwas versteckt, sie besteht darin, dass man die Ziffernfolge, die das Verkaufsprodukt bezeichnet, um eine sogenannte *Prüfziffer* verlängert.

Wir wollen uns hier zuerst der Darstellung der Ziffern durch Striche zuwenden. Der Code beginnt und endet je mit zwei schwarzen Strichen, und auch in der Mitte sind zwei schwarze Striche zu finden. Diese Zeichen dienen der Kalibrierung des Systems. Sie enthalten sonst keine Information. Der ganze Code zerfällt dann in eine linke und eine rechte Seite mit je sechs Ziffern. Die dreizehnte Ziffer ganz links spielt eine Sonderrolle, sie wurde von den Europäern eingeführt; im amerikanischen System ist die Sonderziffer stets 0, wobei diese 0 in Amerika i.a. nicht ausgedruckt wird.

Die Codierung durch Striche

Jede der 12 *normalen* Ziffer wird beim EAN-Code durch ein Strichmuster dargestellt, und zwar sind es je 7 weisse bzw. schwarze Striche gleicher Breite, die eine Ziffer darstellen. Diese Striche werden *ohne Zwischenraum* nebeneinander gedruckt. Wenn somit zwei schwarze oder weisse Striche aufeinanderfolgen, so entsteht ein schwarzer bzw. weisser Strich doppelter Breite, etc. Zur besseren Übersicht symbolisieren wir einen schwarzen Strich durch eine 1 und einen weissen Strich durch eine 0. So gesehen

¹Der Entscheid für den *Universal Product Code*, *UPC* in den USA wurde nach einem Wettbewerb der *National Association of Food Chains* im April 1973 gefällt. Der siegreiche Vorschlag stammte von IBM. In einer gegenüber dem eingereichten Vorschlag leicht veränderten Form wurde dann der UPC-Strichcode in den USA ab Mitte 1974 im Markt eingeführt. Dieser UPC-Code wurde durch den EAN-Code erweitert.

²Diese 13 Ziffern bezeichnen das Produkt eindeutig. Der Aufbau des Ziffernsystems soll hier nicht im Detail betrachtet werden. Bestandteile sind ein "Ländercode" am Anfang, eine Unternehmensnummer und eine Artikelnummer. Die ersten beiden Teilnummern werden zentral verwaltet.

wird somit jede Ziffer durch 7 Bits dargestellt, durch ein *binäres Wort* der Länge 7. Konzentrieren wir uns zuerst auf die rechte Seite. Für diese *rechte Seite* des EAN-Codes verwendet man das folgende Schema:

Set C
Right-Hand Characters
 Even Parity (2 or 4 dark modules)

Module Location:	1	2	3	4	5	6	7
	Binary Pattern (0 or 1 Encodation)						
	↓ ↓ ↓ ↓ ↓ ↓ ↓						
Right-Hand Character of Code:							
0	→ 1	1	1	0	0	1	0
1	→ 1	1	0	0	1	1	0
2	→ 1	1	0	1	1	0	0
3	→ 1	0	0	0	0	1	0
4	→ 1	0	1	1	1	0	0
5	→ 1	0	0	1	1	1	0
6	→ 1	0	1	0	0	0	0
7	→ 1	0	0	0	1	0	0
8	→ 1	0	0	1	0	0	0
9	→ 1	1	1	0	1	0	0

Man sieht, dass hier die Codierung jeder Ziffer mit einer Eins beginnt und mit einer Null endet; ausserdem enthält sie 2 oder 4 Einsen. Die Anzahl Einsen in einem binären Wort nennt man *Parität*. Auf der rechten Seite ist also die Parität immer *gerade*. Hier ist ein Beispiel mit der Ziffer 2 rechts.



Nun zur linken Seite! Wenn wir die *erste* Ziffer der linken Seite des EAN-Codes ansehen, so wird dort offenbar ein anderes Schema verwendet, nämlich das folgende:

Set A
Left-Hand Characters
 Odd Parity (3 or 5 dark modules)

Module Location:	1	2	3	4	5	6	7	
	Binary Pattern (0 or 1 Encodation)							
	↓	↓	↓	↓	↓	↓	↓	
Left-Hand Character of Code:								
0	→	0	0	0	1	1	0	1
1	→	0	0	1	1	0	0	1
2	→	0	0	1	0	0	1	1
3	→	0	1	1	1	1	0	1
4	→	0	1	0	0	0	1	1
5	→	0	1	1	0	0	0	1
6	→	0	1	0	1	1	1	1
7	→	0	1	1	1	0	1	1
8	→	0	1	1	0	1	1	1
9	→	0	0	0	1	0	1	1

Gegenüber dem obigen Schema sind die Ziffern 0 und 1, also weiss und schwarz miteinander vertauscht. Damit beginnt die Codierung der ersten Ziffer links immer mit einer Null und endet mit einer Eins. Ferner weist die Codierung immer 3 oder 5 Einsen auf. Die Parität der ersten Ziffer links ist immer *ungerade*. In unserem Beispiel wird dies bestätigt.



Die unterschiedliche Art der Codierung der äussersten Ziffer, genauer die unterschiedliche Parität – rechts gerade, links ungerade – hat zur Konsequenz, dass das Lesegerät die linke und die rechte Seite des Strichcodes unterscheiden kann. Damit spielt es keine Rolle, ob die Person an der Kasse den EAN-Streifen von links oder von rechts über das Lesegerät bewegt. Darüber haben Sie sich vielleicht schon gewundert.

Man könnte sich vorstellen, dass alle sechs Ziffern auf der linken Seite auf die gleiche Art codiert werden wie die erste. Das ist im amerikanischen System tatsächlich der Fall. Im europäischen System muss allerdings zusätzlich die ganz links stehende Sonderziffer irgendwie codiert werden. Dafür hat man einen recht raffinierten Weg gefunden. Man verwendet für die einzelnen Ziffern auf der linken Seite in Abhängigkeit der Sonderziffer zwei verschiedene Codierungen, nämlich neben derjenigen mit ungerader Parität (siehe oben) auch eine mit gerader Parität. Die Details der Verteilung folgen weiter unten. Hier zuerst die entsprechende Tabelle dieser Codierung:

Set B
Left-Hand Characters
 Even Parity (2 or 4 dark modules)

		Module Location: 1 2 3 4 5 6 7						
		Binary Pattern (0 or 1 Encodation)						
Left-Hand Character of Code:		↓ ↓ ↓ ↓ ↓ ↓ ↓						
0	→	0	1	0	0	1	1	1
1	→	0	1	1	0	0	1	1
2	→	0	0	1	1	0	1	1
3	→	0	1	0	0	0	0	1
4	→	0	0	1	1	1	0	1
5	→	0	1	1	1	0	0	1
6	→	0	0	0	0	1	0	1
7	→	0	0	1	0	0	0	1
8	→	0	0	0	1	0	0	1
9	→	0	0	1	0	1	1	1

Die Codierung jeder Ziffer beginnt hier mit einer Null und endet mit einer Eins und die Ziffern enthalten 2 oder 4 Einsen.

Hier ist noch einmal eine übersichtliche Zusammenfassung der drei Codierungen:

DIGIT	LEFT-HAND ENCODING		RIGHT-HAND ENCODING
	ODD PARITY (A)	EVEN PARITY (B)	ALL CHARACTERS
0	0001101	0100111	1110010
1	0011001	0110011	1100110
2	0010011	0011011	1101100
3	0111101	0100001	1000010
4	0100011	0011101	1011100
5	0110001	0111001	1001110
6	0101111	0000101	1010000
7	0111011	0010001	1000100
8	0110111	0001001	1001000
9	0001011	0010111	1110100

Wie bereits angedeutet werden die zwei unterschiedlichen Codierungen auf der linken Seite verwendet, um die freistehende Sonderziffer darzustellen. Wir haben bereits festgestellt, dass die erste Ziffer links immer mit ungerader Parität codiert wird. Die Parität der Codierungen der anderen 5 Ziffern hängt von der Sonderziffer ab und zwar gemäss dem folgenden Schema:

FIRST NUMBER SYSTEM DIGIT	SECOND NUMBER SYSTEM DIGIT	PARITY TO ENCODE WITH				
		MANUFACTURER CODE CHARACTERS				
		1	2	3	4	5
0 (UPC-A)	Odd	Odd	Odd	Odd	Odd	Odd
1	Odd	Odd	Even	Odd	Even	Even
2	Odd	Odd	Even	Even	Odd	Even
3	Odd	Odd	Even	Even	Even	Odd
4	Odd	Even	Odd	Odd	Even	Even
5	Odd	Even	Even	Odd	Odd	Even
6	Odd	Even	Even	Even	Odd	Odd
7	Odd	Even	Odd	Even	Odd	Even
8	Odd	Even	Odd	Even	Even	Odd
9	Odd	Even	Even	Odd	Even	Odd

Im konkreten Beispiel mit der Sonderziffer 9 wird gemäss Schema für die fünfte Ziffer links, also die 8, die gerade Codierung verwendet.



Im amerikanischen System ist die Sonderziffer immer 0; in diesem Fall werden somit *alle* Ziffern links mit ungerader Parität codiert.

Hinweisen kann man hier noch auf einige weitere Besonderheiten der Codierungen. Man hat darauf geachtet, dass sich die Codierung zweier verschiedener Ziffern an mindestens *zwei* verschiedenen Stellen unterscheiden. In einem gewissen Sinn hat man den “Abstand” zwischen den binären Codierungen zweier Ziffern möglichst gross gemacht. Damit wird der Lesevorgang sicherer und weniger fehleranfällig.

Achtet man darauf, wieviele Gruppen von Einsen die Codierung einer Ziffer aufweist, so sieht man, dass in jeder der drei unterschiedlichen Codierungen immer nur zwei derartige Gruppen auftreten. Innerhalb der Codierung einer Ziffer kommen damit nur höchstens zwei (schmale oder breite) schwarze Balken vor. Auch diese Eigenschaft kann zur Absicherung des Lesevorgangs benützt werden.

Schliesslich zeigt eine genauere Betrachtung der drei Codierungen, dass sie unter sich eng verwandt sind: Die Codierung ungerader Parität auf der linken Seite entsteht aus der Codierung auf der rechten Seite, indem 0 und 1 also weiss und schwarz vertauscht werden. Ferner ist die Codierung gerader Parität auf der linken Seite einfach spiegelverkehrt zur Codierung auf der rechten Seite. Diese Eigenschaft hat seinerzeit sicher die Programmierung des Lesegerätes vereinfacht.

Soweit zur Codierung der Ziffern durch Striche. Ich komme jetzt zum zweiten Grundlage des Strichcodes, zur Prüfziffer.

Die Prüfziffer

Die Produktnummer, von der EAN ausgeht, besteht aus 12 Ziffern, a_1, a_2, \dots, a_{12} . (Wir lesen von links *inklusive* der oben besprochenen freistehenden Sonderziffer!) Diesen 12 Ziffern wird – ganz rechts – eine 13. Ziffer a_{13} angefügt:

$$a_1, a_2, \dots, a_{12}, a_{13} .$$

Und zwar so, dass

$$a_1 + 3a_2 + a_3 + 3a_4 + \cdots + 3a_{12} + a_{13}$$

eine Zehnerzahl ist. Es ist also nicht jede Folge von 13 Ziffern eine EAN-Zahl; dies ist nur dann der Fall, wenn die zusätzliche Bedingung erfüllt ist. Man nennt deshalb a_{13} auch etwa eine *Prüfziffer*. Sie ist ein Sicherheitselement, das Lesefehler unwahrscheinlich machen soll.

An dem folgenden Beispiel sehen wir, wie diese Bedingung zu erfüllen ist.

9	7	8	0	3	8	7	9	4	8	2	3	?
1	3	1	3	1	3	1	3	1	3	1	3	1
9	21	8	0	3	24	7	27	4	24	2	9	?
9	1	8	0	3	4	7	7	4	4	2	9	?
9	1	8	0	3	4	7	7	4	4	2	9	2

Die erste Zeile besteht aus den 12 ersten Ziffern des Codes, die zweite gibt die Gewichte an, mit denen die einzelnen Ziffern zu multiplizieren sind, in der dritten stehen die entsprechenden Produkte. Für die Rechnung modulo 10 müssen Vielfache von 10 nicht berücksichtigt werden, deshalb kann man sich in der vierten Zeile mit den *Resten modulo 10* begnügen. Die Prüfziffer ergänzt die Summe zur nächsten Zehnerzahl.

Liest das Lesegerät die dreizehn Ziffern der EAN-Etikette, so prüft es anschliessend, ob die obige Bedingung erfüllt ist. Ist die Bedingung nicht erfüllt, so ist die gelesene 13stellige Zahl *keine* EAN Nummer, und es muss beim Lesevorgang ein Fehler geschehen sein. Möglicherweise wurde eine oder sogar mehrere Ziffern falsch gelesen, d.h. die entsprechende Strichkombination wurde falsch interpretiert. Der Lesefehler wurde *entdeckt!* In diesem Fall muss der Lesevorgang wiederholt werden.

Die Frage stellt sich: Was genau nützt das Hinzufügen der Prüfziffer? Man wird ja nicht einfach ohne Grund eine weitere, eigentlich "nicht notwendige" Ziffer zur Produktnummer hinzufügen.

Behauptung: Der EAN Code kann einfache Fehler (falsche Ziffern) erkennen.

Beweis: Für ungerade Indizes ist das klar. Deshalb betrachten wir hier nur den Fall von geraden Indizes.

$$a_1 + 3a_2 + a_3 + \cdots + 3a_i + \cdots + a_{13} = x \cdot 10 ,$$

$$a_1 + 3a_2 + a_3 + \cdots + 3b_i + \cdots + a_{13} = y \cdot 10 ,$$

wo x und y ganze Zahlen bezeichnen. Die Differenz ist

$$3a_i - 3b_i = z \cdot 10 ,$$

wo z wiederum eine ganze Zahl ist. Daraus folgt

$$3(a_i - b_i) = z \cdot 10 .$$

Wegen $0 \leq |a_i - b_i| \leq 9$ ist dies nur möglich, wenn $z = 0$, also $a_i = b_i$ ist.

Behauptung: EAN entdeckt das Vertauschen von aufeinanderfolgenden Ziffern, es sei denn, ihre Differenz ist 5.

Beim elektronischen Lesegerät wird dieser Vertauschungs-Fehler wohl kaum auftreten; aber es sind Situationen denkbar, in denen der EAN-Code eingetippt werden muss. Dann sind Sicherheitsmassnahmen gegen das Vertauschen von (aufeinanderfolgenden) Ziffern sinnvoll.

Beweis:

$$a_1 + 3a_2 + \cdots + 3a_i + a_{i+1} + \cdots + a_{13} = x \cdot 10 ,$$

$$a_1 + 3a_2 + \cdots + 3a_{i+1} + a_i + \cdots + a_{13} = y \cdot 10 ,$$

wo x und y ganze Zahlen bezeichnen. Zu zeigen ist: $a_i = a_{i+1}$ oder $|a_i - a_{i+1}| = 5$. Die Differenz der beiden obenstehenden Gleichungen ist:

$$3a_i + a_{i+1} - 3a_{i+1} - a_i = z \cdot 10 ,$$

wo z wiederum eine ganze Zahl ist. Daraus folgt:

$$a_i - a_{i+1} = z \cdot 5 .$$

Wegen $0 \leq |a_i - a_{i+1}| \leq 9$ ist dies nur möglich, wenn $a_i = a_{i+1}$ oder wenn $|a_i - a_{i+1}| = 5$.

Analoges wie für die Vertauschung aufeinanderfolgender Ziffern gilt offenbar für die Vertauschung einer Ziffer an einer geraden Stelle mit einer Ziffer an einer ungeraden Stelle. Im Unterschied dazu kann EAN den Fehler *nicht* erkennen, wenn zwei Stellen mit nur geraden oder nur ungeraden Indizes vertauscht werden.

Das Hinzufügen der Prüfziffer erlaubt es somit, einige Lesefehler zu *entdecken*, andere Fehler bleiben aber trotz der Prüfziffer unentdeckt.

In einer auf der englischen Sprache beruhenden Untersuchung wurde die Fehlerhäufigkeit bei der Übermittlung von Ziffernfolgen (mündlich bzw. durch Eintippen) festgestellt:

Falsche Ziffern 79.1% der Fehler;

Transposition aufeinanderfolgenden Ziffern 10.2% der Fehler;

alle anderen Fehlertypen je unter 1%, darunter der häufigste Fehler mit 0.8% Vertauschung von *abc* zu *bca*.

Bei der Kommunikation über die Sprachgrenzen hinweg, also zwischen deutsch und englisch oder zwischen deutsch und französisch dürfte der Fehlertypus der Transposition einen wesentlich höheren Anteil haben. Denken Sie etwa an “twentyfive” und “fünfundzwanzig”.

Innerhalb des Gesamtsystems des EAN-Codes, also einschliesslich des zugehörigen Lesegerätes, kann man sich eigentlich nur Lesefehler vorstellen, die durch die Prüfwert entdeckt werden. Die Fehlerquote des Systems dürfte damit sehr nahe bei Null liegen. Erst wenn die Ziffernreihe des EAN-Codes eingetippt werden muss, sind vereinzelte Fehler denkbar.

Weitere Bemerkungen

Der EAN-Strichcode hat sich seit seiner Einführung gegen Ende der 70er Jahre flächendeckend durchgesetzt. Es gibt auch heute, fast 40 Jahre nach der Einführung, meines Wissens kein Detailhandelsgeschäft, das diesen Strichcode nicht verwendet. So gesehen ist der EAN-Strichcode ein Muster-Beispiel guter Ingenieurarbeit.

Über den Detailhandel hinaus hat der Strichcode viele weitere Anwendungen gefunden, oft in einer etwas abgeänderten Form. Beispiele sind: Der ETH-Ausweis, die Bezeichnung der Bücher der ETH-Bibliothek (und auch anderer Bibliotheken), die Cumulus-Karte der Migros, die Supercard von Coop, etc.

Auch die neue AHV-Nummer verwendet die EAN-Codierung mit der Prüfwert. Dabei hat man eine seltsame Teil-Konstruktion gewählt. Zwar enthält die neue AHV-Nummer die EAN-Prüfwert, aber auf den zugehörigen Strichcode hat man verzichtet. Ich befürchte, dass dies längerfristig zu Problemen führt, denn eine Vertauschung von zwei aufeinanderfolgenden Ziffern, wie sie bei der mündlichen Übermittlung bzw. beim Eintippen oft vorkommt, wird – wie wir gesehen haben – nicht in allen Fällen entdeckt! Ich jedenfalls würde mein AHV-Konto sehr sorgfältig im Auge behalten, wenn meine AHV-Nummer zwei aufeinanderfolgende Ziffern enthält, die sich um 5 unterscheiden.

Zur Illustration möge das Folgende dienen.³ Man stelle sich vor, dass das Büro des Arbeitgebers sich bei der Überweisung des AHV-Beitrages vertippt und in der AHV-Nummer die beiden Ziffern 9 und 4 vertauscht. Die untenstehende Rechnung zeigt, dass die falsche EAN-Nummer die gleiche Prüfziffer aufweist wie die richtige. Die Vertauschung wird also nicht entdeckt und die Einzahlung landet auf dem falschen Konto!

9	7	8	0	3	8	7	9	4	8	2	3	2
9	7	8	0	3	8	7	4	9	8	2	3	?
1	3	1	3	1	3	1	3	1	3	1	3	1
9	21	8	0	3	24	7	12	9	24	2	9	?
9	1	8	0	3	4	7	7	2	9	2	9	2
9	7	8	0	3	8	7	4	9	8	2	3	2

Ziffernfolgen werden übrigens heute fast überall durch Prüfziffern abgesichert; natürlich nicht unbedingt durch die Prüfziffer des EAN-Codes. Denn neben der EAN-Prüfziffer gibt es anders konstruierte, darunter ganz raffinierte und viel leistungsfähigere. Ein Beispiel, das bereits vor dem EAN-Code verwendet worden ist, ist die ISBN-Nummer des Buchhandels (siehe die obere Zahlenreihe auf dem reproduzierten EAN-Code). Erst später hat der Buchhandel auf den EAN-Code umgestellt. Weitere Beispiele von Prüfziffern treten bei der Nummer der Kreditkarte auf, bei der IBAN-Nummer für Banküberweisungen, bei der Nummer des PC-Kontos, bei der Codierzeile des Posteingahlungsscheins, etc. Auch gewisse Sicherheitselemente auf Banknoten basieren auf diesem Prinzip; hier sind die Grundlagen natürlich geheim. Die Liste liesse sich beliebig verlängern.

³Der Einfachheit halber nehmen wir zur Illustration die EAN-Nummer unserer EAN-Etikette; sie ist natürlich keine echte AHV-Nummer.

technischen Möglichkeiten viel zu hoch gewesen wäre. Ich denke, dass heute die technischen Voraussetzungen vorhanden wären, um ohne Prüzfiffer auszukommen. Aber das ist *nicht* der Lauf der Dinge: Die technischen Möglichkeiten werden viel mehr dazu eingesetzt, um kompliziertere Strich- und andere Codes zu verwenden, die auf kleinerem Raum mehr Information als gewöhnliche EAN-Etiketten enthalten. So werden z.B. zur fabrikinternen Identifikation der Produkte oft viel dichtere und längere Strichcodes verwendet. Ein konkretes Beispiel sehen Sie hier; es handelt sich um die Verpackung eines Notebooks.



Ferner gibt es heute Codes, welche die Information nicht nur eindimensional sondern zweidimensional speichern. Beispiele sind die elektronischen Briefmarken der Schweizer und der Deutschen Post (Matrix-Code).





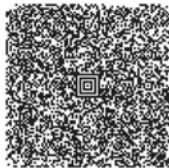
Dann aber auch die elektronischen Billete der SBB und der DB (Aztec-Code).

SBB CFF FFS OT 5774 9472 2177

Gewöhnliche Billette Stammach-Stockler 1

GÜLTIG: 2015		15.05.1941	
30	⌚	VON/DE/DA → NACH/A	30
28.05	*	ZÜRICH HB	→ NESSLAU-NEU ST. JOHANN
*	*	*	→ *
VIA RAPPERSWIL-WATTWIL *****			
1/2		CHF 14.50	


Bestellnummer: 52387462 ARTIKEL-NR: 125
INKL. 08.00% MWST/SBB




Für das OnlineTicket gelten spezielle Konditionen gemäss den «Allgemeinen Geschäftsbedingungen» unter www.sbb.ch/ticketshop.

Auszug:

- Die auf dem OnlineTicket erwähnte Person muss dem Kontrollpersonal unaufgefordert den A4-Ausdruck des OnlineTickets, einen amtlichen Ausweis (Reisepass oder Identitätskarte) sowie allfällige Ermässigungskarten (z.B. Halbtax-Abonnement, BahnCard usw.) vorweisen.
- Reisende, die in der Schweiz ein OnlineTicket mit Halbtax-Ermässigung haben, weisen anstelle eines amtlichen Ausweises das gültige Halbtax-Abonnement vor.
- Ausdruck: 100% Grösse (nicht skaliert), weiss, zuvor unbedrucktes A4-Papier, Hochformat, Auflösung mind. 600 dpi. Die Werbung im unteren Drittel darf abgetrennt werden.
- Das Online Ticket ist nicht übertragbar.
- Kein Umtausch, keine Erstattung (Ausnahme: Für internationale Fahrweise gelten teilweise andere Bestimmungen).





26050926 © SBB AG 1052.11

Von Japan herkommend haben sich neuerdings die sogenannten QR-Codes stark verbreitet; QR steht für **Q**uick **R**esponse. Solche Codes finden sich häufig auf Plakaten und in Inseraten. Sie können von Smartphones “photographiert” werden, worauf man automatisch Kontakt zur Internetseite erhält, auf die das Inserat bzw. das Plakat hinweist. Die QR-Matrix enthält üblicherweise die Internetadresse, das Smartphone ist in der Lage, diese zu lesen und die entsprechende Seite auf Knopfdruck via Internet zu kontaktieren. Konkret findet man solche QR-Codes zum Beispiel auf dem

DB-Reiseplan, in Zeitungsinseraten, sogar in Jobinseraten, etc. Einer Pressemeldung entnehme ich, dass QR-Codes auch für das Nachfolgesystem der Einzahlungsscheine verwendet werden sollen.

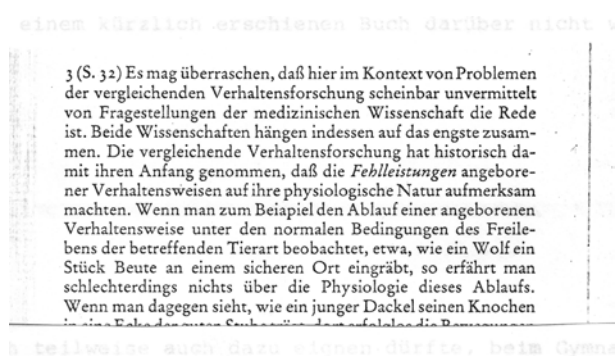


Das alles wäre kaum möglich, wenn dieses QR-System nicht einen sehr leistungsfähigen *fehlerkorrigierenden Code* verwenden würde. Damit sind wir auf einem höheren Level als der Fehlererkennung angelangt, nämlich bei der *Fehlerkorrektur*. Codes, nicht nur die QR-Codes, sondern z.B. auch derjenige, der bei den SBB-Fahrkarten verwendet wird, müssen sehr sicher sein, und die darin vorkommende Information muss auch dann noch zweifelsfrei gelesen werden können, wenn die Codes leicht beschädigt sind. Vielleicht wurde z.B. das Papier gefaltet, sodass der Falz gerade über das Code-Quadrat verläuft. Deshalb verwenden diese Codes nicht nur einfache Prüfziffern. Sie sind so konstruiert, dass sie in gewissen Grenzen die *Korrektur* von Fehlern ermöglichen.

Das ist ein faszinierendes Gebiet; man stelle sich vor, da kann ein technisches System entstandene Fehler selbständig korrigieren. So etwas hätte man sich bei Vorlesungen gewünscht, nämlich, dass die Studierenden Fehler automatisch korrigieren würden!

Natürlich kann ich hier nicht auf die teilweise sehr mathematischen Grundlagen eingehen, aber ich möchte mit einigen Bemerkungen zeigen, wie eine solche Korrektur *prinzipiell* vonstatten geht. Dazu zuerst ein Beispiel eines Druckfehlers *Beiapiel* in einem gewöhnlichen Text. Dieser lässt sich bei jedem Lesen sofort korrigieren; es

ist jedermann klar, dass das *Beispiel* heissen sollte. Die Fehlerkorrektur ist möglich! Weshalb? Wir nehmen automatisch an, dass das Druckfehlerwort “nahe” bei einem “richtigen” Wort liegt, dass also nur *ein* Buchstabe falsch ist. Prinzipiell könnte der Druckfehler aber auch auf eine andere Weise zustande gekommen sein, nämlich so, dass *a* richtig ist, aber die anderen oder einige der anderen Buchstaben falsch sind! Aber das Auftreten eines solchen Fehlers wäre viel unwahrscheinlicher, und wir verwerfen beim Lesen diese Möglichkeit ganz automatisch.



Genau diesem Prinzip folgt die *technische* Fehlerkorrektur. Dies sei im Folgenden kurz erklärt. Wir beschränken uns dabei auf 01-Codes. Da sorgt man dafür, dass der Code nur eine beschränkte Zahl von “richtigen” Wörter enthält. Fehler lassen dann i.a. ein Wort entstehen, das *nicht* zum Code gehört. Der Fehler wird erstens erkannt und dann nach maximaler Wahrscheinlichkeit zum *nächstliegenden* Wort korrigiert.

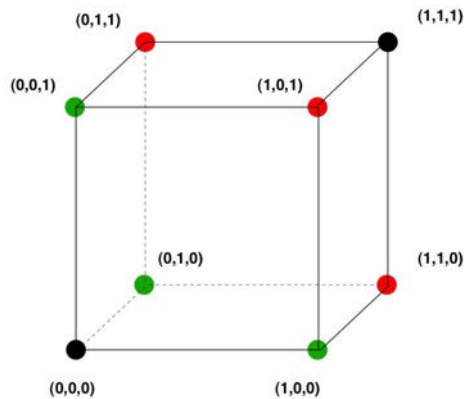
Beispiel **Repetitionscode (dreifach)**:

Aus 010... mache man 000111000...; man wiederhole jedes Bit dreimal.

Liest der Empfänger 000101000..., so wird er erstens feststellen, dass beim zweiten Nachrichtenbit ein Fehler eingetreten ist, und er wird zweitens aus Wahrscheinlichkeitsgründen annehmen, dass an zweiter Stelle des Nachrichtenwortes eigentlich eine 1 stehen sollte.

Der Empfänger kann einfache Fehler nicht nur entdecken, sondern auch korrigieren.

Ich mache ein Beispiel, ein triviales(!) Beispiel: Man benütze einen Code, der jedes Bit verdreifacht. Zulässige Wörter des Codes sind also nur 000 und 111. Wenn nun im Laufe der Übermittlung ein Fehler geschieht, etwa 101, so wird dies bei der Fehlerkorrektur als 111 gelesen. Weshalb? Weil es wahrscheinlicher ist, dass nur an *einer* Stelle ein Fehler geschehen ist als gleichzeitig an zwei Stellen.








Die für die Fehlerkorrektur verwendeten Codes sind natürlich viel raffinierter, als der hier nur als Beispiel verwendete *dreifache Repetitionscode*. Die “richtigen” Codes verwenden viel Mathematik und lassen sich nicht in wenigen Worten beschreiben. Aber genau wie Prüfziffern werden solche Codes heute an fast allen Stellen verwendet, an denen grosse digitale Datenmengen im Spiel sind. Davon merkt der Benutzer i.a kaum etwas.

Ein schönes konkretes Beispiel ist die CD. Die Compact Disk speichert die Musik digital. Zu diesem Zweck wird das Audiosignal auf jedem der zwei Kanäle 44'100 mal pro Sekunde abgetastet und in eines der $2^{16} = 65'536$ Niveaus eingeteilt. Jede Messung liefert also 16 Bits. Man kann sich leicht ausrechnen, wie gross die Datenmenge für eine Musikaufnahme ist, die eine Stunde dauert. Es ist klar, dass eine derart lange 01-Folge sich mit einem normalen Aufwand nicht fehlerfrei abspeichern und ablesen lässt.⁴ Schon aus diesem Grund muss bei diesem System eine automatische

⁴In Tat und Wahrheit ist die abzuspeichernde Datenmenge aus verschiedenen Gründen sogar noch wesentlich grösser, nämlich etwa dreimal so gross. Obschon die Aufzeichnung der einzelnen Bits mikroskopisch klein ist, ist die Tonspur auf einer normalen CD sehr lang, sie hat eine Länge von etwa 6 km.

Fehlerkorrektur eingebaut werden. Das ist in der Tat der Fall, es sind in der CD sogar zwei sehr leistungsfähige fehlerkorrigierende Codes eingebaut. Damit erreicht man auch andere positive Effekte; so können etwa Ablesestörungen, die ihren Grund in Staubpartikeln auf der CD haben, ebenfalls korrigiert werden. Ja, sogar massive Kratzer auf der CD beeinträchtigen die Qualität der Wiedergabe im allgemeinen nicht. Wenn man sich an die alten Schallplatten erinnert, so ist dies eine erhebliche Verbesserung!

Ähnliche fehlerkorrigierende Codes wie bei der CD werden bei deren Weiterentwicklung, der DVD eingesetzt, dann auch bei der Audio-Aufzeichnungssoftware MP3, der Bildsoftware JPEG, bei der digitalen Verbreitung von Radio- und Fernsehprogrammen, bei den Telefonverbindungen via Handy, etc.⁵ Und eben auch bei dem oben erwähnten QR-Code. Beim QR-Code gibt es offenbar verschiedene Korrekturlevels: für den stärksten Level ist nach den offiziellen Angaben eine Korrektur der Daten sogar dann möglich, wenn die Matrix zu 30% beschädigt ist. Damit sind der Verwendung dieses Codes kaum Grenzen gesetzt.

<p>Sie erinnern sich: Musik will etwas ausdrücken, will etwas an das Publikum übermitteln. Das können sehr viele verschiedene Dinge sein. Musik kann etwa einfach gute Laune verströmen ...</p> <p>(Scott Joplin (1867/1868–1917): <i>Maple Leaf Rag</i> (Beginn))⁸</p> <p>oder auch schlechte Laune ...</p> <p>(Robert Schumann (1810–1856): <i>Pantalon et Colombine</i> (Beginn) aus <i>Carnaval</i> op. 9)⁹</p> <p>Musik kann todtraurig sein ...</p> <p>(Franz Schubert (1797–1828): <i>Andantino</i> (Beginn) aus der Sonate in A-Dur, D 959)¹⁰</p> <p>oder himmelhochjauchzend ...</p> <p>(Franz Schubert (1797–1828): <i>Impromptu As-Dur, D 899, Nr. 4</i> (Schluss))¹¹</p> <p>Musik kann Eleganz ausstrahlen, und was kann das besser als ein Walzer von Chopin?</p> <p>(Frédéric Chopin (1810–1849): <i>Grande Valse Brillante Es-Dur, op. 18</i> (Beginn))¹²</p> <p>Wir kommen zu <i>Humor in der Musik</i>. Das ist für sich ein abendfüllendes Thema. Der Großmeister des Humors in der Musik war ohne Zweifel Joseph Haydn. Sie kennen alle seinen berühmtesten Scherz: jenen Orchesterschlag in der – wie wir auf Deutsch sagen – „Symphonie mit dem Paukenschlag“, oder – wie die Englischsprachigen ein wenig treffender sagen: der <i>Surprise Symphony</i>,</p>	    
---	--

Zwei Beispiele dazu: Kürzlich wurde in einer Zeitschrift der Text eines Vortrages über

⁵Zur Übermittlung von Audio- und Fernsehprogrammen ist vielleicht noch anzumerken, dass es gerade die Digitalisierung ist, welche eine Fehlerkorrektur überhaupt erlaubt. Bei einer analogen Übertragung lassen sich einmal eingetretene Störungen bzw. Verluste nicht oder allenfalls nur mit grösstem Aufwand rückgängig machen. Das ist einer der Gründe – nicht der einzige –, weshalb man neuerdings bei Radio und Fernsehen die digitale Übermittlung bevorzugt.

Musik veröffentlicht. Der Vortrag beinhaltete natürlich auch Musikbeispiele, die in einer Textveröffentlichung nicht so ohne weiteres einzubinden sind. Ein Ausweg bot der QR-Code an: zu einzelnen Textstelle wurde ein Codequadrat mitgedruckt, das direkt per URL auf entsprechende Musikstück im Internet wies. Per Smartphone lässt sich also der ganze Vortrag mit aller Musik miterleben.

Hier ist ein weiteres Beispiel einer originellen Verwendung des QR-Codes; wenn Sie wollen, können Sie dieses sogar als eine mögliche Handlungsanweisung auffassen: es ist - dank der robusten Fehlerkorrekturmöglichkeiten - ohne weiteres möglich, ein QR-Code-Quadrat mit der URL Ihrer eigenen Homepage auf einem T-Shirt anzubringen. Wenn Sie das tragen, sind Sie für die ganze Umwelt dank Smartphone völlig transparent.



Vielleicht noch ein Wort zur Zukunft. Es scheint, dass die EAN-Etikette bald einmal abgelöst wird, sehr wahrscheinlich durch einen RFID-Tag. (RFID steht für **R**adio **F**requency **I**dentification.) Damit erfolgt die Identifikation berührungslos. In gewissen Bereichen ist dieses System bereits in Betrieb. So funktioniert das Ausleihsystem der Zentralbibliothek seit einiger Zeit auf diese Weise.

Wenn man im Detailhandel dieses System einführen sollte, so wird man natürlich auch für die Kreditkarte des Kunden ein System verwenden, das auf dem gleichen Prinzip beruht. In der Tat gibt es schon Stellen, wo diese Bezahlweise verwendet wird. Beispielsweise ist dies bei den Automaten der ZVV der Fall.

Der Ersatz des EAN-Strichcodes durch RFID wird natürlich Begleiteffekte haben. Der Käufer wird noch "durchsichtiger" als er es bis anhin schon war; und es ist denkbar, dass der RFID-Tag der Kreditkarte nicht nur beim Bezahlvorgang abgelesen wird, sondern auch an anderen Stellen. Es ist mindestens denkbar, dass auf diese Weise sogar ein vollständiges Bewegungsprofil des Trägers erstellt werden kann.

Berührungsloses Einscannen der eingekauften Gegenstände wie auch das berührungslose Bezahlen sind an sich angenehme Entwicklungen. Meine Sorge ist, dass diese Systeme auch mehr und mehr eingesetzt werden, um die Personen und ihr Verhalten auszuhorchen und zu überwachen. Hoffen wir, dass unsere Gesellschaft diese offensichtlich unangenehme, ja gefährliche Seite der Entwicklung unter Kontrolle halten kann.

Oktober 2015

Prof. em. Urs Stammbach