Prof. Dr. Alfred Toth

## Zu einer semiotischen Codierungstheorie

1. Im folgenden publiziere ich ein undatiertes Typoskript, das ich möglicherweise noch in den 1980er Jahren, da ich als Mathematiker bei Max Bense ein Zweitstudium absolviert hatte, vor allem aber unter seinem Einfluß in Zusammenhang mit der bekanntlich von Bense stark beeinflußten Informationstheorie (vgl. etwa das Vorwort Georg Heikes zur 2. Aufl. von Meyer-Epplers Klassiker [Meyer-Eppler 1969, S. v], das mit Max Benses Werk beginnt) geschrieben hatte. Das Werk ist zwar, was seinen Gegenstand, die Grundlagen einer semiotischen Codierungstheorie, betrifft, vollständig, bildet aber trotzdem nur ein Kapitel eines nie geschriebenen Buches. Das Manuskript, in dem die Definitionen aus einer mir nicht mehr eruierbaren Quelle offenbar wörtlich entommen wurden (vgl. zur Thematik Schulz 1991), ist erst gestern zufällig wieder aufgetaucht, und ich publiziere es hiermit, indem ich die 7 Seiten photomechanisch reproduziere.

| Nr. | Zkl | | | Zeilenvektoren | Spaltenvektoren |
|---|---|---|---|---|---|
| 1 | 3.1 | 2.1 | 1.1 | 1 0 0 1 0 0 1 0 0 | 1 1 1 0 0 0 0 0 0 |
| 2 | 3.1 | 2.1 | 1.2 | 0 1 0 1 0 0 1 0 0 | 0 1 1 1 0 0 0 0 0 |
| 3 | 3.1 | 2.1 | 1.3 | 0 0 1 1 0 0 1 0 0 | 0 1 1 0 0 0 1 0 0 |
| 4 | 3.1 | 2.2 | 1.2 | 0 1 0 0 1 0 1 0 0 | 0 0 1 1 1 0 0 0 0 |
| 5 | 3.1 | 2.2 | 1.3 | 0 0 1 0 1 0 1 0 0 | 0 0 1 0 1 0 1 0 0 |
| 6 | 3.1 | 2.3 | 1.3 | 0 0 1 0 0 1 1 0 0 | 0 0 1 0 0 0 1 1 0 |
| 7 | 3.2 | 2.2 | 1.2 | 0 1 0 0 1 0 0 1 0 | 0 0 0 1 1 1 0 0 0 |
| 8 | 3.2 | 2.2 | 1.3 | 0 0 1 0 1 0 0 1 0 | 0 0 0 0 1 1 1 0 0 |
| 9 | 3.2 | 2.3 | 1.3 | 0 0 1 0 0 1 0 1 0 | 0 0 0 0 0 1 1 1 0 |
| 10 | 3.3 | 2.3 | 1.3 | 0 0 1 0 0 1 0 0 1 | 0 0 0 0 0 0 1 1 1 |

**Vollständige Generatormatrix (Zeilenvektoren)**

| | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 |
| 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 0 |
| 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 0 |
| 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 0 |
| 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 |

**Vollständige Generatormatrix (Spaltenvektoren)**

| | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 |
| 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 |

**Vollständige Kontrollmatrix (Zeilenvektoren)**

$$
\begin{matrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{matrix}
$$

**Vollständige Kontrollmatrix (Spaltenvektoren)**

$$
\begin{matrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{matrix}
$$

Die **Codewörter** sind die Zeilenvektoren $X = (x_1, \ldots, x_n)$. Für alle Codewörter gilt die Beziehung

(1) $P X^T = 0$.

Diese Beziehung, welche die Grundlage für die Feststellung von in Codewörtern aufgetretenen Fehlern ist, kann auch in der Form

(2) $x_1 P_1 + x_2 P_2 + x_3 P_3 + \ldots + x_n P_n = 0$

notiert werden. Die Vektoren $P_i$ ($i = 1, \ldots, n$) heißen **Prüfvektoren**. Diese sind die Spaltenvektoren der Kontrollmatrix.

Nun wird angenommen, daß ein Codewort $X = (x_1, x_2, \ldots, x_n)$ gefälscht wird und stattdessen eine Kombination $X' = (x'_1, x'_2, \ldots, x'_n)$ empfangen wird. Dann ist dieser Fehler – wenn überhaupt – dadurch erkennbar, daß (1) bzw. (2) nicht mehr erfüllt ist, also

(3) $$Z = x'_1\,P_1 + x'_2\,P_2 + x'_3\,P_3 + \ldots + x'_n\,P_n \neq 0$$

Z ist – ebenso wie die Prüfvektoren – ein Spaltenvektor mit k Elementen, man nennt Z den **Fehlervektor** oder auch **Syndrom**. $Z = 0$ bedeutet, daß kein (erkennbarer) Fehler aufgetreten ist, ansonsten kann aus der Form von Z auf die Art des Fehlers geschlossen werden.

Das fehlerhafte Muster kann auch in der Form

(4) $$X' = X + F$$

dargestellt werden. Dabei ist $F = (f_1, f_2, f_3, \ldots f_n)$ ein aus n Stellen bestehender Zeilenvektor, der an den gestörten Stellen von X eine 1 aufweist, an den anderen eine 0.

**Beispiel**

Als Codewort wählen wir das 2. Generatorwort aus der Matrix:

$$X^z = (0 \quad 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1\ 0\ 1\ 0\ 0\ 1\ 0\ 0) \quad \text{bzw.}$$
$$X^s = (0 \quad 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1\ 1\ 1\ 0\ 0\ 0\ 0\ 0)$$

Gestört werden die Stellen 3 und 7, so daß

$$X'^z = (0 \quad 1\ 1\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 1\ 0\ 1\ 0\ 0\ 1\ 0\ 0) \quad \text{bzw.}$$
$$X'^s = (0 \quad 1\ 1\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 1\ 1\ 1\ 0\ 0\ 0\ 0\ 0)$$

empfangen wird. Das Fehlermuster hat dann in der 3. und 7. Stelle eine 1:

$$F = (0 \quad 0\ 1\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0).$$

Man erkennt, daß die modulo 2-Addition X + F auf das gefälschte Muster X' führt.

Aus (4) folgt

(5) $X' = X + F = (x'_1, x'_2, x'_3, \ldots, x'_n) = (x_1 + f_1, x_2 + f_2, x_3 + f_3, \ldots, x_n + f_n)$

und somit hat T nach (3) auch die Form

(6) $$Z = (x_1 + f_1)P_1 + (x_2 + f_2)P_2 + (x_3 + f_3)P_3 + \ldots + (x_n + f_n)P_n = (x_1P_1 + x_2P_2 + \ldots x_nP_n) + (f_1P_1 + f_2P_2 + \ldots + f_nP_n).$$

Die ersten durch die Klammer (2. Zeile) zusammengefaßten Summanden stellen nach (1) die Prüfbedingungen für ein nicht gestörtes Codewort dar; diese Summe verschwindet also. Wir erhalten damit

(7) $Z = f_1P_1 + f_2P_2 + f_3P_3 + \ldots + f_nP_n.$

Dieses Ergebnis ist bemerkenswert, weil zur Fehlerprüfung das richtige Codewort X selbst ohne Bedeutung ist, lediglich die auftretenden Fehlermuster $F = (f_1, f_2, \ldots, f_n)$ spielen eine Rolle. Offenbar können genau die Fehlermuster F erkannt werden, bei denen (7) auf einen Vektor $Z \neq 0$ führt.

Wir prüfen nun, welche Bedingungen an die Prüfvektoren zu stellen sind, damit ein Fehler, zwei Fehler usw. erkannt werden können. Diese Frage entspricht der Ermitlung der Hamming-Distanz des Codes.

a) Ein aufgetretener Fehler
Wenn ein Fehler im Codewort auftritt, hat genau eine Stelle $f_i$ des Fehlermusters den Wert 1, alle anderen den Wert 0. Nach (7) erhalten wir dann

$$Z = f_iP_i = P_i, \; i = 1 \ldots n.$$

die Bedingung $Z \neq 0$ tritt auf, wenn $P_i \neq 0$ ist. Falls keiner der n Prüfvektoren verschwindet, ist ein auftretender Fehler in dem Codewort erkennbar. Der Nullvektor ist also als Prüfvektor nicht zugelassen.

b) Zwei aufgetretene Fehler
Im Fehlermuster haben zwei Stellen $f_i$, $f_j$ ($i \neq j$) den Wert 1. (7) liefert

$$Z = f_iP_i + f_jP_j = P_i + P_j, \; i, j = 1 \ldots n, \; i \neq j.$$

Die Fehler werden erkannt, wenn $Z \neq 0$ ist, und das heißt, daß die Summe zweier beliebiger Prüfvektoren nicht verschwinden darf. Anders gesagt bedeutet dies, daß zwei beliebige Prüfvektoren linear voneinander unabhängig sein müssen. Lineare Unabhängigkeit liegt genau dann vor, wenn sich alle Prüfvektoren voneinander unterscheiden. Dabei darf der Nullvektor kein Prüfvektor sein.

c) Drei aufgetretene Fehler
Entsprechende Überlegungen wie oben führen zu der Bedingung

$$Z = P_i + P_j + P_k = 0, \; i \neq j \neq k.$$

Dieses bedeutet, daß beliebige 3er-Kombinationen von Prüfvektoren linear voneinander unabhängig sein müssen.

Wie sieht es mit den Vollständigen semiotischen Kontrollmatrizen aus? Keiner der 19 Prüfvektoren (Spalten der Matrix P) ist der Nullvektor; der Code kann also mindestens einen Fehler erkennen. Da sich alle Prüvektoren voneinander unterscheiden, sind je zwei Prüfvektoren linear unabhängig. Dies gilt ebenfalls für 3er-Kombinationen, denn es scheint keine Gleichung der Form $P_i + P_j = P_k$ (mit $i \neq j \neq k$) erfüllt zu sein. Es gibt jedoch Beispiele für die Gleichung der Form $P_i + P_j + P_k = P_l$ (mit $i \neq j \neq k \neq l$):

$$P_5 + P_6 + P_8 = P_9.$$

Da also der semiotische Code 3 Fehler erkennen kann, hat er eine Hamming-Distanz von $h = 4$.

Ein allgemeiner Weg zur Lokalisierung einer falschen Stelle eines Codes besteht darin, daß zu jedem möglichen Fehlermuster der Fehlervektor Z (das Syndrom) berechnet wird und aus diesem auf den Fehler rückgeschlossen wird.

a) Fehler in der 1. Codewortstelle:

$F = (1\ 0\ \ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0)$

Nach (7) und den Prüfvektoren erhält man

$Z = P_1 = (1\ \ \ 0\ 0\ 1\ 0\ 0\ 1\ 0\ 0)^T$ (= 3.1 2.1 1.1)

b) Fehler in der 1. und 5. Codewortstelle:

F = (1 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0)

$Z = P_1 + P_5 = (1\ \ 0\ 0\ 1\ 0\ 0\ 1\ 0\ 0)^T + (0\ \ 0\ \ 1\ 0\ 1\ 0\ 1\ 0\ 0)^T =$
$(1\ 0\ \ 1\ 1\ 1\ 0\ 0\ 0\ 0)^T$ (= 2.2 2.1 1.3 1.1)

c) Fehler in der 1., 5. und 9 Codewortstelle:

F = (1 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0)

$Z = P_1 + P_5 + P_9 = (1\ \ 0\ 0\ 1\ 0\ 0\ 1\ 0\ 0)^T + (0\ \ 0\ \ 1\ 0\ \ 1\ 0\ 1\ 0\ 0)^T +$
$(0\ \ 0\ 1\ 0\ 0\ 1\ 0\ 1\ 0)^T = (1\ \ 0\ 0\ 1\ 1\ 1\ 0\ 1\ 0)^T$
(= 3.2 2.3 2.2 2.1 1.1)

**Syndrom-Decodierung:**

Sei $C \subseteq V$ ein linearer Code, der t-fehlerkorrigierend ist. Man erstellt eine Liste der nebenklassenanführer (= Fehlermuster) und der zugehörigen Syndrome:

| Nebenklassenanführer | Syndrome |
|---|---|
| 1 0 0 0 0 0 0 0 0 | 1 0 0 1 0 0 1 0 0 |
| 0 1 0 0 0 0 0 0 0 | 0 1 0 1 0 0 1 0 0 |
| 0 0 1 0 0 0 0 0 0 | 0 0 1 1 0 0 1 0 0 |
| 0 0 0 1 0 0 0 0 0 | 0 1 0 0 1 0 1 0 0 |
| 0 0 0 0 1 0 0 0 0 | 0 0 1 0 1 0 1 0 0 |
| 0 0 0 0 0 1 0 0 0 | 0 0 1 0 0 1 1 0 0 |
| 0 0 0 0 0 0 1 0 0 | 0 1 0 0 1 0 0 1 0 |
| 0 0 0 0 0 0 0 1 0 | 0 0 1 0 1 0 0 1 0 |
| 0 0 0 0 0 0 0 0 1 | 0 0 1 0 0 1 0 1 0 |
| 0 0 0 0 0 0 0 0 0 | 0 0 1 0 0 1 0 0 1 |

Für einen empfangenen Vektor x

1. Berechnet man das Syndrom s(x)

Sei x = (1 0 1 0 1 0 0 1 0). In Frage kommen folgende Syndrome:

7

| | |
|---|---|
| wegen 1 an 1. Stelle: | (1 0 0 1 0 0 1 0 0) |
| wegen 1 an 3. Stelle: | (0 0 1 1 0 0 1 0 0) |
| wegen 1 an 5. Stelle: | (0 0 1 0 1 0 1 0 0) |
| wegen 1 an 8. Stelle: | (0 0 1 0 1 0 0 1 0) |
| modulo 2-Addition: | (1 0 1 0 0 0 1 1 0) |

2. Sucht dies in der Liste der Syndrome:

   existiert nicht!

3. Stellt den zugehörigen Nebenklassenanführer fest:

   unmöglich!

4. Decodiert x zu c := x + e

   unmöglich!

Literatur


Meyer-Eppler, W[olfgang], Grundlagen und Anwendungen der Informationstheorie. 2. Aufl. Berlin 1969

Schulz, Ralph-Hardo, Codierungstheorie. Braunschweig 1991


8.1.2016