Informationstheorie – eine mathematische Theorie der Datenkompression und Datenübertragung

Erhard Aichinger

Institut für Algebra
Johannes Kepler Universität Linz

JKU Mathematikseminar 2016

Ziele der Informations- und Codierungstheorie

Claude Elwood Shannon (1916 - 2001)

The fundamental problem of communciation is that of reproducing at one point either exactly or approximately a message selected at another point.

Ziele der Informations- und Codierungstheorie

Claude Elwood Shannon (1916 - 2001)

The fundamental problem of communciation is that of reproducing at one point either exactly or approximately a message selected at another point.

Ziele der Informations- und Codierungstheorie

Ziele der Informations- und Codierungstheorie

Claude Elwood Shannon (1916 - 2001)

The fundamental problem of communciation is that of reproducing at one point either exactly or approximately a message selected at another point.

Ziele der Informations- und Codierungstheorie

- Effizientes Darstellen einer Nachricht (Datenkompression)

Ziele der Informations- und Codierungstheorie

Claude Elwood Shannon (1916 - 2001)

The fundamental problem of communciation is that of reproducing at one point either exactly or approximately a message selected at another point.

Ziele der Informations- und Codierungstheorie

- Effizientes Darstellen einer Nachricht (Datenkompression)
- Zeitsparendes Schicken der Nachricht über einen Nachrichtenkanal

Ziele der Informations- und Codierungstheorie

Claude Elwood Shannon (1916 - 2001)

The fundamental problem of communciation is that of reproducing at one point either exactly or approximately a message selected at another point.

Ziele der Informations- und Codierungstheorie

- Effizientes Darstellen einer Nachricht (Datenkompression)
- Zeitsparendes Schicken der Nachricht über einen Nachrichtenkanal
- Korrektur von Übertragungsfehlern

Ein Kommunikationssystem

Modell eines Kommunikationssystems

Ein Kommunikationssystem

Modell eines Kommunikationssystems

- Nachrichtenquelle

Ein Kommunikationssystem

Modell eines Kommunikationssystems

- Nachrichtenquelle
- Quellcodierer

Ein Kommunikationssystem

Modell eines Kommunikationssystems

- Nachrichtenquelle
- Quellcodierer
- Kanalcodierer

Ein Kommunikationssystem

Modell eines Kommunikationssystems

- Nachrichtenquelle
- Quellcodierer
- Kanalcodierer
- Kanal und Rauschen

Ein Kommunikationssystem

Modell eines Kommunikationssystems

- Nachrichtenquelle
- Quellcodierer
- Kanalcodierer
- Kanal und Rauschen
- Kanaldecodierer

Ein Kommunikationssystem

Modell eines Kommunikationssystems

- Nachrichtenquelle
- Quellcodierer
- Kanalcodierer
- Kanal und Rauschen
- Kanaldecodierer
- Quelldecodierer

Ein Kommunikationssystem

Modell eines Kommunikationssystems

- Nachrichtenquelle
- Quellcodierer
- Kanalcodierer
- Kanal und Rauschen
- Kanaldecodierer
- Quelldecodierer
- Nachrichtensenke

Quellcodierung

Eine typische Nachricht:

```
AABAAAAAAAAACAAAAAAAABAAAAAAAAAEAAAAAAAAAAAEAAEAAEAAAAA
AAAEAACAABBAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
EAAAAAAAAAAAAAAEAAAAAABAAAAABAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAEAAEAAAAAAAAAAAAAAAAABAAAAAAAAAAAAAAAAAAAAB
EAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACAAAACAAAAAAAAAABAAAAAAB
AAAAAAAAAAAAAAAAAAAACABAAAABAAAAEAAAAAAAAAAAAABAAAAAAAA
CAAAABAAAAAAEAAAABAAAAAAAAAAAACAAAAAAAAAAAAAAAABAAAAAAA
BAAAAAACAAABAAAAAAAAAAAAAAAAAAAAAABAAAAAAAAAAAAAAAAAAAB
AAAAAAAAABAAAAAABAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAABAAAABB
AAAAA
```

Woher kommt diese typische Nachricht?

An jeder Stelle der Nachricht tritt mit Wahrscheinlichkeit 0.9 ein `A`, mit 0.06 ein `B`, mit 0.015 ein `C`, mit 0.015 ein `D` und mit Wahrscheinlichkeit 0.01 ein `E` auf.

Woher kommt diese typische Nachricht?

An jeder Stelle der Nachricht tritt mit Wahrscheinlichkeit 0.9 ein `A`, mit 0.06 ein `B`, mit 0.015 ein `C`, mit 0.015 ein `D` und mit Wahrscheinlichkeit 0.01 ein `E` auf.

Ziele

Woher kommt diese typische Nachricht?

An jeder Stelle der Nachricht tritt mit Wahrscheinlichkeit 0.9 ein `A`, mit 0.06 ein `B`, mit 0.015 ein `C`, mit 0.015 ein `D` und mit Wahrscheinlichkeit 0.01 ein `E` auf.

Ziele

1. Eine solche Nachricht als Folge aus 0 und 1 codieren.

Woher kommt diese typische Nachricht?

An jeder Stelle der Nachricht tritt mit Wahrscheinlichkeit 0.9 ein `A`, mit 0.06 ein `B`, mit 0.015 ein `C`, mit 0.015 ein `D` und mit Wahrscheinlichkeit 0.01 ein `E` auf.

Ziele

1. Eine solche Nachricht als Folge aus 0 und 1 codieren.
2. Im Durchschnitt nur höchstens 0.8 Bits / Zeichen benötigen – eine Folge von 100 Zeichen sollte im Durchschnitt also auf 80 Bits komprimiert werden können.

Quellcodierung

Ein anderes Beispiel

Nachricht:

$AAABAACAADBBAAABAAAAABAAAAAAABAADAAAA\ldots$.

Quellcodierung

Ein anderes Beispiel

Nachricht:
$AAABAACAADBBAAABAAAAABAAAAAAAABAADAAAA \ldots.$

- 60%A, 30%B, 5%C, 5%D.

Quellcodierung

Ein anderes Beispiel

Nachricht:
$AAABAACAADBBAAABAAAAABAAAAAAABAADAAAA \ldots .$

- $60\%A$, $30\%B$, $5\%C$, $5\%D$.

Zeichenweise Codierung der Nachrichten als $0/1$-Folgen

Zeichen	Vorschlag 1	Vorschlag 2	Vorschlag 3
A	00	0	0
B	01	10	10
C	10	110	110
D	11	01	111
	2	oje	1.5

Quellcodierung

Ziel der Quellcodierung

Ziel der Quellcodierung ist die Datenkompression.

Quellcodierung

Shannons Quellcodierungssatz - I

Wenn die Zeichen $A_1, A_2, \ldots, A_n$ unabhängig voneinander mit Wahrscheinlichkeiten $p_1, p_2, \ldots, p_n$ auftreten, so braucht jedes Quellcodierungsverfahren, das für beliebig lange Dateien funktioniert, im Mittel zumindest

$$\sum_{i=1}^{n} p_i \cdot \log_2\left(\frac{1}{p_i}\right)$$

Bits pro Nachrichtenzeichen. Durch geeignete Codierungsverfahren kann man dieser Schranke beliebig nahe kommen.

Quellcodierung

Shannons Quellcodierungssatz - I

Wenn die Zeichen $A_1, A_2, \ldots, A_n$ unabhängig voneinander mit Wahrscheinlichkeiten $p_1, p_2, \ldots, p_n$ auftreten, so braucht jedes Quellcodierungsverfahren, das für beliebig lange Dateien funktioniert, im Mittel zumindest

$$\sum_{i=1}^{n} p_i \cdot \log_2(\frac{1}{p_i})$$

Bits pro Nachrichtenzeichen. Durch geeignete Codierungsverfahren kann man dieser Schranke beliebig nahe kommen.

Für das obige Beispiel ist diese Schranke ungefähr 1.395 Bits pro Nachrichtenzeichen.

Die beste zeichenweise Codierung

Huffman-Algorithmus [Ash, 1990, p. 42], [MacKay, 2003, p.99]

Für gegebene Zeichen $A_1, A_2, \ldots, A_n$ mit Wahrscheinlichkeiten $p_1, p_2, \ldots, p_n$ produziert der Huffman-Algorithmus die beste zeichenweise Codierung von $A_1, A_2, \ldots, A_n$ als $0/1$-Folgen.

Kanalcodierung

Binärer symmetrischer Kanal

Der binäre symmetrische Kanal überträgt pro Durchgang ein Bit. Jedes Bit wird mit Wahrscheinlichkeit $1 - f$ richtig übertragen. (f ist die Fehlerwahrscheinlichkeit des Kanals).

Beispiel für Kanalkodierung

Beispiel für Kanalkodierung

Beispiel für Kanalkodierung

Beispiel für Kanalkodierung

1. Wir wollen die Bitfolge $x_1 x_2 x_3 x_4 \ldots$ übertragen.

Beispiel für Kanalkodierung

Beispiel für Kanalkodierung

1. Wir wollen die Bitfolge $x_1 x_2 x_3 x_4 \ldots$ übertragen.
2. Der Kanalcodierer fügt Kontrollstellen dazu und transformiert $x_1 x_2 x_3 x_4 \ldots$ zu $y_1 y_2 y_3 y_4 y_5 y_6 \ldots$.

Beispiel für Kanalkodierung

Beispiel für Kanalkodierung

1. Wir wollen die Bitfolge $x_1 x_2 x_3 x_4 \dots$ übertragen.
2. Der Kanalcodierer fügt Kontrollstellen dazu und transformiert $x_1 x_2 x_3 x_4 \dots$ zu $y_1 y_2 y_3 y_4 y_5 y_6 \dots$.
3. $y_1 y_2 y_3 y_4 y_5 y_6 \dots$ wird über den Kanal gesendet. Die Folge $z_1 z_2 z_3 z_4 z_5 z_6 \dots$ kommt an.

Beispiel für Kanalkodierung

Beispiel für Kanalkodierung

1. Wir wollen die Bitfolge $x_1 x_2 x_3 x_4 \ldots$ übertragen.
2. Der Kanalcodierer fügt Kontrollstellen dazu und transformiert $x_1 x_2 x_3 x_4 \ldots$ zu $y_1 y_2 y_3 y_4 y_5 y_6 \ldots$.
3. $y_1 y_2 y_3 y_4 y_5 y_6 \ldots$ wird über den Kanal gesendet. Die Folge $z_1 z_2 z_3 z_4 z_5 z_6 \ldots$ kommt an.
4. Der Kanaldecodierer versucht jene Folge $x_1 x_2 x_3 x_4 \ldots$ zu finden, die am wahrscheinlichsten gesendet wurde. Er produziert eine Folge $u_1 u_2 u_3 u_4 \ldots$.

Bitfehlerrate und Übertragungsrate

Definition (Bitfehlerrate)

Die mittlere Bitfehlerrate b dieser Prozedur ist die Wahrscheinlichkeit dafür, dass $x_i \neq u_i$.

Bitfehlerrate und Übertragungsrate

Definition (Bitfehlerrate)

Die mittlere Bitfehlerrate b dieser Prozedur ist die Wahrscheinlichkeit dafür, dass $x_i \neq u_i$.

Definition (Übertragungsrate)

Die Übertragungsrate r ist die Anzahl der Quellbits pro gesendetem Kanalbit.

Shannons Kanalkodierungssatz

Die Frage

Wir haben einen Kanal mit Fehlerrate f, und eine Übertragungsrate r vorgegeben. Welche Bitfehlerrate können wir bei Verwendung dieses Kanals mit Übertragungsrate r bestenfalls erreichen?

Shannons Kanalcodierungssatz

Theorem (Shannons Kanalcodierungssatz)

Sei C der binäre symmetrische Kanal mit Fehlerrate f.

Shannons Kanalcodierungssatz

Theorem (Shannons Kanalcodierungssatz)

Sei C der binäre symmetrische Kanal mit Fehlerrate f.

1. Wenn die Übertragungsrate r die Ungleichung

$$r < 1 + f \cdot \log_2(f) + (1 - f) \cdot \log_2(1 - f),$$

   erfüllt, und $\varepsilon > 0$ gilt, dann gibt es eine Übertragungsprozedur mit Übertragungsrate r, sodass die Bitfehlerrate $< \varepsilon$ ist.

Shannons Kanalcodierungssatz

Theorem (Shannons Kanalcodierungssatz)

Sei C der binäre symmetrische Kanal mit Fehlerrate f.

1. Wenn die Übertragungsrate r die Ungleichung

$$r < 1 + f \cdot \log_2(f) + (1 - f) \cdot \log_2(1 - f),$$

   erfüllt, und $\varepsilon > 0$ gilt, dann gibt es eine Übertragungsprozedur mit Übertragungsrate r, sodass die Bitfehlerrate $< \varepsilon$ ist.
2. Wenn $r > 1 + f \cdot \log_2(f) + (1 - f) \cdot \log_2(1 - f)$, dann gibt es ein $b > 0$, sodass die Bitfehlerrate jedes Übertragungssystems mit Rate r zumindest b ist.

Ash, R. B. (1990).
Information theory.
Dover Publications Inc., New York.
Corrected reprint of the 1965 original.

Ash, R. B. (1990).
Information theory.
Dover Publications Inc., New York.
Corrected reprint of the 1965 original.

MacKay, D. J. C. (2003).
Information theory, inference and learning algorithms.
Cambridge University Press, New York.
The book can be viewed at
http://www.inference.phy.cam.ac.uk/mackay/itprnn